

Lucas de Assis Soares

**Segmentação de Bacilos de Tuberculose em
Imagens de Microscopia Convencional Através
da Utilização de Redes Neurais Artificiais e
Máquinas de Vetores Suporte**

Vitória - ES, Brasil

2015

Lucas de Assis Soares

Segmentação de Bacilos de Tuberculose em Imagens de Microscopia Convencional Através da Utilização de Redes Neurais Artificiais e Máquinas de Vetores Suporte

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Engenharia Elétrica.

Universidade Federal do Espírito Santo – UFES

Departamento de Engenharia Elétrica

Programa de Pós-Graduação em Engenharia Elétrica

Orientador: Klaus Fabian Côco

Co-orientador: Evandro Ottoni Teatini Salles

Vitória - ES, Brasil

2015

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica,
Universidade Federal do Espírito Santo, ES, Brasil)

S676s Soares, Lucas de Assis, 1989-
Segmentação de bacilos de tuberculose em imagens de
microscopia convencional através da utilização de redes neurais
artificiais e máquinas de vetores suporte / Lucas de Assis
Soares. – 2015.
83 f. : il.

Orientador: Evandro Ottoni Teatini Salles.
Coorientador: Klaus Fabian Côco.
Dissertação (Mestrado em Engenharia Elétrica) –
Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Processamento de imagens. 2. Reconhecimento de
padrões. 3. Tuberculose – Diagnóstico. 4. Corantes e coloração
(Microscopia). 5. Microscopia. I. Salles, Evandro Ottoni Teatini.
II. Côco, Klaus Fabian. III. Universidade Federal do Espírito
Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

Agradecimentos

Aos meus orientadores Prof. Dr. Klaus Fabian Côco e Prof. Dr. Evandro Ottoni Teatini Salles pela confiança em mim depositada e a oportunidade de desenvolver este trabalho;

À minha família, pelo apoio incondicional em todas minhas decisões;

À minha namorada, por sempre tornar meus dias melhores;

À Deus, pelo dom da vida e da inteligência;

E a todos os amigos que de alguma forma contribuíram para este trabalho.

Resumo

Embora tratável, a tuberculose (TB) é um problema de saúde global estando apenas atrás da AIDS como o maior responsável por mortes devido a um único agente infeccioso (WHO, 2015). Para que seja tratada, um diagnóstico adequado da doença deve ser feito. Ainda hoje, a maior parte do diagnóstico laboratorial é realizado usando a técnica de coloração de Ziehl-Neelsen sobre a lâmina com escarro do paciente, e esta é examinada por um microscópio ótico em busca dos bacilos de tuberculose através da inspeção humana. Esse processo é demorado e exaustivo, de modo que um sistema de reconhecimento automático de bacilos de tuberculose, a partir de imagens da lâmina, permitiria tornar o processo de diagnóstico mais ágil e menos cansativo. Neste trabalho, um sistema de segmentação automática de bacilos de tuberculose usando imagens de microscopia convencional é proposto. O sistema é basicamente dividido em duas etapas: a segmentação propriamente dita e a classificação das estruturas segmentadas. Primeiramente, as imagens são projetadas com base na análise do discriminante linear de Fisher, permitindo uma maior separação entre os *pixels* de fundo e os *pixels* de bacilo. Em seguida, duas abordagens são testadas: uma segmentação baseada em limiarização global e outra baseada no método de Otsu. As estruturas muito grandes e muito pequenas passam por um processo de filtragem e é feito um pós-processamento morfológico. Por fim, a classificação das estruturas segmentadas é feita usando redes neurais artificiais e máquinas de vetores suporte. Os resultados mostram que é possível implementar um sistema de segmentação automática de bacilos de tuberculose com boa capacidade de distinção dos bacilos e de baixo custo computacional. Para a segmentação, até 98,69% dos bacilos são corretamente segmentados e até 85,61% dos bacilos permanecem após o filtro de áreas. Para a classificação das estruturas, foram obtidos valores médios de sensibilidade, especificidade, acurácia e precisão de até 94,25%, 95,33%, 95,73% e 92,50%, respectivamente.

Palavras-chaves: Segmentação. Processamento de imagens. Reconhecimento de padrões. Baciloscopia óptica. Ziehl-Neelsen. Microscopia convencional. Bacilos de tuberculose.

Abstract

Even though tuberculosis (TB) is a treatable disease, it is still a major global health problem being second only to AIDS as the greatest killer worldwide due to a single infectious agent (WHO, 2015). In order to treat it, the disease must be properly diagnosed. The diagnosis is usually done using by staining a slide with patient sputum using the Ziehl-Neelsen stain and then a human specialist analyzes it using an optical microscope looking for tuberculosis bacilli. Since this process is time consuming and labor intensive, an automatic bacilli recognition system allows the diagnosis process to be more agile and less tiresome. In this work, an automatic tuberculosis bacilli segmentation system using conventional microscopy images is proposed. The system is basically divided in two stages: a stage of segmentation and a stage for classification of the segmented structures. First, images are projected based on a linear discriminant analysis considering Fisher criterion in order to increase the separation between bacilli pixels and background pixels. Then, two approaches are evaluated: a segmentation process based on global thresholding and another based on Otsu segmentation method. Structures that have a big or a small area are then filtered and morphological operators are applied over the binary image. Finally, the segmented structures are classified using artificial neural networks and support vector machines. The results show that it is possible to implement an automatic tuberculosis bacilli segmentation system that provides a good distinction of bacilli in the images with a low computational cost. For the segmentation stage, up to 98.69% of bacilli are correctly segmented and up to 85.51% of bacilli remain after the area filter. For the classification of the structures, mean values up to 94.25%, 95.33%, 95.73% and 92.50% were obtained for sensitivity, specificity, accuracy and precision, respectively.

Key-words: Segmentation. Image processing. Pattern recognition. Optical bacilloscopy. Ziehl-Neelsen. Conventional microscopy. Tuberculosis bacilli.

Lista de ilustrações

Figura 1 – Diagrama de blocos do sistema desenvolvido.	22
Figura 2 – Elipse, com destaque para os eixos maior e menor. Fonte: Stewart (2012)	28
Figura 3 – Modelo de um neurônio. Fonte: (HAYKIN et al., 2009)	31
Figura 4 – Modelo alternativo de um neurônio. Fonte: (HAYKIN et al., 2009) . . .	32
Figura 5 – Rede neural alimentada diretamente com uma única camada oculta. Fonte: (HAYKIN et al., 2009)	32
Figura 6 – Vetores suporte em um problema de classificação envolvendo duas classes linearmente separáveis. Fonte: (ALPAYDIN, 2004)	35
Figura 7 – Histograma das áreas de bacilos.	38
Figura 8 – Exemplos de imagens do banco de dados.	39
Figura 9 – Exemplos de imagens do banco de dados marcadas pelo especialista. . .	40
Figura 10 – Distribuição de cores conforme o matiz.	41
Figura 11 – Nível de cinza utilizado para a escolha do limiar.	42
Figura 12 – Relação entre o limiar escolhido e a intensidade média das imagens projetadas.	43
Figura 13 – Função de primeiro grau obtido por regressão linear da relação entre o limiar escolhido e a intensidade média das imagens projetadas.	43
Figura 14 – Histograma das áreas de bacilos.	44
Figura 15 – Elemento estruturante na forma de um disco com raio de 2 <i>pixels</i>	45
Figura 16 – Imagem gerada pela segmentação baseada no método de Otsu.	47
Figura 17 – Projeção das imagens mostradas na Figura 8.	50
Figura 18 – Segmentação das imagens mostradas na Figura 17 pelo método de limiarização global.	51
Figura 19 – Resultado da eliminação de áreas muito grandes e muito pequenas e pós-processamento morfológico das imagens mostradas na Figura 18. . .	52
Figura 20 – Segmentação das imagens mostradas na Figura 17 pelo Método de Otsu modificado.	53
Figura 21 – Resultado da eliminação de áreas muito grandes e muito pequenas e pós-processamento morfológico das imagens mostradas na Figura 20. . .	54
Figura 22 – Histograma da projeção das imagens mostradas na Figura 17.	58
Figura 23 – Algumas das imagens relativas ao sexto paciente.	61

Lista de tabelas

Tabela 1 – Resumo dos valores de média e desvio padrão para cada <i>fold</i> da classificação das estruturas segmentadas usando redes neurais.	55
Tabela 2 – Resumo dos valores de média e desvio padrão para cada <i>fold</i> da classificação das estruturas segmentadas usando máquinas de vetores suporte.	56
Tabela 3 – Resumo das médias de cada <i>fold</i> da classificação usando redes neurais.	59
Tabela 4 – Resumo das médias de cada <i>fold</i> da classificação usando máquinas de vetores suporte.	59
Tabela 5 – Média e desvio padrão da taxa de classificação de estruturas duvidosas como bacilos usando redes neurais artificiais e máquinas de vetores suporte como classificadores.	60
Tabela 6 – Comparação entre a classificação de estruturas com o trabalho realizado por Costa Filho et al. (2012).	62
Tabela 7 – Sensibilidade devido à classificação por redes neurais.	72
Tabela 8 – Especificidade devido à classificação por redes neurais.	73
Tabela 9 – Acurácia devido à classificação por redes neurais.	74
Tabela 10 – Precisão devido à classificação por redes neurais.	75
Tabela 11 – Resumo dos valores de média e desvio padrão por paciente para a classificação das estruturas segmentadas usando redes neurais artificiais.	76
Tabela 12 – Taxa de classificação de estruturas duvidosas como bacilos usando redes neurais.	76
Tabela 13 – Sensibilidade devido à classificação por máquinas de vetores suporte.	78
Tabela 14 – Especificidade devido à classificação por máquinas de vetores suporte.	79
Tabela 15 – Acurácia devido à classificação por máquinas de vetores suporte.	80
Tabela 16 – Precisão devido à classificação por máquinas de vetores suporte.	81
Tabela 17 – Resumo dos valores de média e desvio padrão por paciente para a classificação das estruturas segmentadas usando máquinas de vetores suporte.	82
Tabela 18 – Taxa de classificação de estruturas duvidosas como bacilos usando máquinas de vetores suporte.	82

Lista de abreviaturas e siglas

BAAR	Bacilos Álcool-Ácido Resistentes
CDC	<i>Centers for Disease Control and Prevention</i>
CR	<i>Color Ratio</i>
FN	<i>False negatives</i> - Falsos negativos
FP	<i>False positives</i> - Falsos positivos
HIV	<i>Human immunodeficiency virus</i> - Vírus da imunodeficiência humana
MS	Ministério da Saúde
OMS	Organização Mundial da Saúde
RNA	Redes Neurais Artificiais
SVM	<i>Support vector machine</i>
TB	Tuberculose
TN	<i>True negatives</i> - Verdadeiros negativos
TP	<i>True positives</i> - Verdadeiros positivos
WHO	<i>World Health Organization</i>

Lista de símbolos

x	Valor escalar
\mathbf{x}	Vetor
\mathbf{X}	Matriz
\mathbf{x}^t	Transposto
J	Razão de discriminante de Fisher
\mathbf{v}	Vetor de projeção obtida pela análise de discriminante linear
χ_i	Classe i
$\boldsymbol{\mu}_i$	Média da classe i
m_i	Média da projeção da classe i
s_i	Espalhamento em torno da média da projeção da classe i
\mathbf{p}	Vetor que representa os três canais de cores de um <i>pixel</i> usado na análise de discriminante linear
u	Projeção do <i>pixel</i> obtida pela análise de discriminante linear
\mathbf{S}_B	Matriz de espalhamento entre-classes
\mathbf{S}_W	Matriz de espalhamento intra-classes
T	Limiar de segmentação
$f(x, y)$	Intensidade de cinza do <i>pixel</i> (x, y) antes da segmentação
$g(x, y)$	Intensidade de cinza do <i>pixel</i> (x, y) após a segmentação
σ^2	Função critério definida no método de Otsu
P_i	Probabilidade de que um <i>pixel</i> seja atribuído à classe i
ν	Média da intensidade de cinza dos <i>pixels</i> de cada classe na aplicação do método de Otsu
q_i	Quantidade de <i>pixels</i> em uma imagem com intensidade de cinza i
L	Quantidade de níveis de cinza existentes em uma imagem

I	Imagem
E	Elemento estruturante
\ominus	Operação morfológica de erosão
\oplus	Operação morfológica de dilatação
\circ	Operação morfológica de abertura
\bullet	Operação morfológica de fechamento
E_x	Excentricidade
b	Comprimento do menor eixo da elipse
a	Comprimento do maior eixo da elipse
C	Circularidade
A	Área de uma estrutura
Q	Perímetro de uma estrutura
ρ_{pq}	Momento de ordem $(p + q)$
κ_{pq}	Momento central de ordem $(p + q)$
η_{pq}	Momento central normalizado de ordem $(p + q)$
ϕ_i	i -ésimo momento invariante
$F(o)$	o -ésimo descritor de Fourier
x_i	Entrada i do classificador
\mathbf{x}	Vetor de entradas do classificador
w_{km}	Peso do neurônio k em relação à entrada x_m
\mathbf{w}_k	Vetor de pesos sinápticos do neurônio k
b_k	<i>Bias</i> do neurônio k
v_k	Potencial de ativação do neurônio k
$\varphi(\cdot)$	Função de ativação do neurônio
y_k	Saída do neurônio k
\mathbf{y}	Vetor de saídas de uma camada de uma rede neural

\mathbf{W}	Matriz de pesos sinápticos
h	Rótulo da entrada quando da utilização de máquinas de vetores suporte
ω	Vetor de projeção usado na classificação usando máquinas de vetores suporte
d	Distância da amostra \mathbf{x} ao hiperplano discriminante
B	Fator de penalização
ξ	Variável que quantifica o erro na classificação
$\gamma(\cdot)$	Função de projeção da amostra \mathbf{x} em um espaço de maior dimensão
\mathbf{z}	Projeção da amostra em um espaço de maior dimensão
Υ	Função de <i>kernel</i>
D	Medida de distância

Sumário

1	INTRODUÇÃO	14
1.1	Objetivos	15
1.1.1	Objetivo geral	15
1.1.2	Objetivos específicos	16
1.2	Revisão da literatura	16
1.3	Estrutura do trabalho	21
2	REFERENCIAIS TEÓRICOS	22
2.1	Análise de discriminante linear	23
2.2	Segmentação das imagens	25
2.2.1	Segmentação por limiarização global	25
2.2.2	Método de Otsu	25
2.3	Operações morfológicas em imagens digitais binárias	26
2.3.1	Operação de abertura	26
2.3.2	Operação de fechamento	26
2.4	Representação de estruturas em imagens digitais	27
2.4.1	Excentricidade	27
2.4.2	Circularidade	27
2.4.3	Momentos invariantes	28
2.4.4	Descritores de Fourier	29
2.4.5	Descritores de cores	30
2.5	Classificadores	30
2.5.1	Redes neurais artificiais	31
2.5.2	Máquinas de vetores suporte	33
3	METODOLOGIA	37
3.1	Imagens utilizadas no trabalho	37
3.2	Sistema proposto	38
3.2.1	Segmentação das imagens	38
3.2.1.1	Segmentação baseada em limiarização global	41
3.2.1.2	Eliminação de áreas	43
3.2.1.3	Pós-processamento morfológico	45
3.2.1.4	Segmentação baseada no método de Otsu	45
3.2.2	Classificação das estruturas	46
4	RESULTADOS	49

4.1	Projeção das imagens	49
4.2	Segmentação das imagens	49
4.2.1	Segmentação baseada em limiarização global	49
4.2.2	Segmentação baseada no método de Otsu	51
4.3	Classificação das estruturas	55
4.3.1	Classificação usando redes neurais	55
4.3.2	Classificação usando máquinas de vetores suporte	55
5	ANÁLISE DOS RESULTADOS	57
5.1	Projeção das imagens	57
5.2	Segmentação das imagens	57
5.3	Classificação das estruturas	59
5.4	Comparação com outros trabalhos	60
6	CONCLUSÃO	63
6.1	Sugestões para trabalhos futuros	63
	Referências	65
	APÊNDICES	68
	APÊNDICE A – DIAGNÓSTICO DE TUBERCULOSE POR BACI- LOSCOPIA	69
A.1	Preparação das lâminas	69
A.2	Leitura e interpretação dos resultados	70
	APÊNDICE B – RESULTADO DA CLASSIFICAÇÃO DAS ESTRU- TURAS SEGMENTADAS	71
B.1	Classificação usando redes neurais	71
B.2	Classificação usando máquinas de vetores suporte	76

1 Introdução

Segundo a Organização Mundial de Saúde (OMS), a tuberculose (TB) é uma doença infecciosa causada pela *Mycobacterium tuberculosis*. Ela usualmente afeta os pulmões, mas também pode atingir outros tecidos como o fígado, a espinha dorsal e o cérebro (WHO, 2015).

A doença é transmitida de pessoa para pessoa por meio de gotículas expelidas pelo tórax e dos pulmões de pessoas infectadas. As bactérias da tuberculose são então colocadas no ar quando pessoas doentes tosse, espirram, falam ou cantam, contaminando pessoas próximas que respiram essas bactérias e se contaminam (CDC, 2014).

A infecção pela *Mycobacterium tuberculosis* normalmente não causa sintomas em pessoas saudáveis devido ao fato de que o sistema imunológico normalmente as defendem das bactérias, tendo a doença na sua forma passiva. Entretanto, se uma pessoa desenvolve tuberculose ativa em seus pulmões, os sintomas são tosse, geralmente acompanhada de escarro ou sangue, dores no peito, fraqueza, perda de peso, febre e sudorese noturna (WHO, 2015).

Segundo a Organização Mundial de Saúde, em 2013, cerca de 9 milhões de pessoas desenvolveram tuberculose e 1,5 milhão de pessoas morreram por causa da doença, sendo 360.000 delas HIV-positivo. A doença só fica atrás da AIDS como a responsável pelo maior número de mortes devido a um único agente infeccioso (WHO, 2014). O número de mortes devido à tuberculose ainda é inaceitavelmente grande devido ao fato de que a doença é prevenível e tratável. Por sua grande ocorrência, a doença ainda permanece como um dos maiores problemas de saúde global (WHO, 2015).

Em 2013, mais de 95% das mortes devido à tuberculose ocorreram em países subdesenvolvidos (WHO, 2014). No Brasil, a cada ano, são notificados aproximadamente 70 mil novos casos e ocorrem 4,6 mil mortes em decorrência da doença. O país ocupa o 17º lugar entre os 22 países responsáveis por 80% do total de casos de tuberculose no mundo (MINISTÉRIO DA SAÚDE, 2015).

A doença é normalmente tratada através de antibióticos e, sem tratamento, a taxa de mortalidade devido à tuberculose é alta. Para tratar a doença de forma apropriada é necessário diagnosticá-la. O método mais comum de diagnóstico laboratorial de tuberculose em países em desenvolvimento é a baciloscopia do escarro, na qual uma amostra de escarro é coletada, colorida e examinada em um microscópio por um técnico de laboratório especializado.

Duas técnicas são comumente utilizadas para o diagnóstico de tuberculose pela

microscopia do escarro: a microscopia de fluorescência e a microscopia convencional. A primeira técnica utiliza um corante fluoróforo, enquanto o segundo utiliza um corante carbol-fucsina e recebe o nome de técnica de Ziehl-Neelsen (UNIVERSITY OF UTAH, 2015).

A microscopia por fluorescência exige uma fonte de luz fluorescente intensa, mas precisa de pouca ampliação em suas lentes objetivas, de modo que a área de um escarro pode ser examinada muito mais rápido em comparação à microscopia convencional. O método também apresenta uma maior sensibilidade do que a técnica de Ziehl-Neelsen. Por outro lado, o sistema de microscopia é mais caro do que o usado em microscopia convencional e exige uma manutenção especializada (COSTA FILHO et al., 2012; DESIKAN, 2013).

Devido aos altos custos associados com a microscopia por fluorescência, o método preferido em países em desenvolvimento é a microscopia convencional usando a técnica de Ziehl-Neelsen. Nesse método, a amostra é coberta com o corante de carbol-fucsina por alguns minutos e depois de aquecida, molhada e tratada com uma solução de álcool-ácido, a amostra é contrastada com outro corante, o azul de metileno, para colorir o fundo da amostra e obter o contraste entre as bactérias e o fundo. Quando seca, os bacilos presentes na lâmina devem aparecer com uma coloração rosa-avermelhada em contraste com o fundo azul (SMART, 2007).

Para fazer o diagnóstico de tuberculose, o técnico de laboratório deve examinar meticulosamente cada lâmina procurando por bacilos e então registrar o número de bactérias observadas na lâmina (SMART, 2007). Esse processo é cansativo, trabalhoso e demorado. O Apêndice A apresenta com detalhes o procedimento de preparação (Seção A.1) e leitura (Seção A.2) de lâminas para o diagnóstico de tuberculose. Por essa razão, é interessante o desenvolvimento de um sistema capaz de analisar a imagem registrada por um microscópio e identificar os bacilos de tuberculose presentes na imagem de forma automática.

Este trabalho pretende desenvolver um sistema de identificação automática de bacilos de tuberculose em imagens de microscopia convencional. **A hipótese, neste caso, é que é possível desenvolver um sistema de identificação automática de bacilos de tuberculose em imagens de microscopia convencional baseado em técnicas de processamento de imagens digitais e reconhecimento de padrões.**

1.1 Objetivos

1.1.1 Objetivo geral

Desenvolver um sistema de identificação automática de bacilos de tuberculose em imagens de microscopia convencional usando técnicas de processamento de imagens digitais

e reconhecimento de padrões.

1.1.2 Objetivos específicos

- Analisar a distribuição de cada um dos canais de cor dos *pixels* das imagens de microscopia convencional;
- Identificar as ferramentas de processamento de imagens digitais que podem ser utilizadas para a segmentação da imagem;
- Desenvolver uma metodologia adequada para a segmentação das imagens coloridas pela técnica de Ziehl-Neelsen;
- Identificar as ferramentas de reconhecimento de padrões e processamento de imagens digitais que podem ser utilizadas para a classificação das estruturas segmentadas;
- Desenvolver uma metodologia adequada para a classificação das estruturas segmentadas;
- Avaliar o sistema desenvolvido, segundo os métodos conhecidos na literatura, comparando-o com outros sistemas desenvolvidos para o mesmo fim.

1.2 Revisão da literatura

Os primeiros trabalhos desenvolvidos no campo de reconhecimento automático de bacilos de tuberculose trataram de imagens de microscopia por fluorescência.

Veropoulos et al. (1998) propõem a utilização de técnicas de processamento de imagens e o uso de redes neurais para a identificação automática de bacilos de tuberculose em imagens de microscopia de fluorescência. A segmentação é feita usando um detector de bordas de Canny seguido da remoção de estruturas devido a seu tamanho. As estruturas são identificadas utilizando os quinze maiores descritores de Fourier. A acurácia média obtida na validação cruzada usando *k-fold* foi de 97,57% acompanhada de uma sensibilidade de 93,53% e uma especificidade de 98,79%.

Forero-Vargas et al. (2001) propõem a segmentação de imagens de baciloscopia provenientes de microscopia de fluorescência usando segmentação baseada em conjuntos *fuzzy*. A classificação das estruturas é feita usando um filtro de correlação de fase (*phase-only correlation filtering*). Nenhuma avaliação quantitativa dos resultados foi realizada.

Forero-Vargas et al. (2002) propõem a segmentação de imagens de baciloscopia provenientes de microscopia de fluorescência usando segmentação baseada em conjuntos *fuzzy* seguido de uma operação morfológica *top-hat* que consiste na subtração da imagem

original pela operação abertura da imagem. A classificação das estruturas é feita utilizando discriminante linear de Fisher. Os parâmetros analisados são área, perímetro, o comprimento do maior e do menor eixo, alongação, circularidade, níveis de cinza máximo e mínimo e densidade do nível de cinza. Os parâmetros definidos como de maior importância para a classificação foram a área, a densidade do nível de cinza, o comprimento do eixo maior e a circularidade. O resultado obtido foi uma especificidade de 82%.

Sobre trabalhos que tratam de microscopia convencional, Sadaphal et al. (2008) propõem a segmentação das imagens baseada em uma análise Bayesiana, onde calcula-se a probabilidade de um pixel representar um bacilo a partir de informação a priori das cores no método de Ziehl-Neelsen. Além disso, propõe-se a dilatação com um elemento estruturante circular para melhorar a imagem segmentada. Para a classificação das estruturas, propõe-se a análise do comprimento do maior eixo e da excentricidade em uma árvore de decisão. Nenhuma avaliação quantitativa dos resultados foi realizada.

Costa et al. (2008) propõem a segmentação na imagem formada pela operação ($R - G$), ou seja, a subtração da componente verde (*green*) da componente vermelha (*red*) da imagem original. A segmentação é feita com limiarização global cujo limiar é definido como a banda do histograma com quantidade de pixels superior a um valor determinado empiricamente. Finalmente, estruturas com menos de 200 pixels são eliminadas. Os melhores resultados alcançados são de 76,65% e 12% para a sensibilidade e a especificidade, respectivamente. No entanto, não são fornecidos detalhes de como esses valores foram alcançados.

Khutlang et al. (2010) propõem a detecção automática de bacilos de tuberculose usando dois classificadores de uma classe. O primeiro classificador é responsável por classificar os pixels a partir das informações de cor. Foram analisados os espaços de cor RGB, HSI, YCbCr e Lab, avaliando classificadores Gaussiano, de mistura de Gaussianas e de análise de componentes principais. Em seguida, são eliminadas as áreas menores que 50 pixels e áreas maiores que 400 pixels. O segundo classificador é responsável por identificar os bacilos a partir das estruturas segmentadas. Foram avaliados a excentricidade, a compacidade, momentos invariantes e descritores de Fourier, e a combinação dessas características usando análise de discriminante linear. Novamente, foram avaliados os classificadores Gaussiano, de mistura de Gaussianas, de análise de componentes principais e k -vizinhos-mais-próximos.

A avaliação do primeiro classificador é feita usando a taxa de pixels corretamente e incorretamente segmentados. O espaço de cores que apresentou melhores resultados foi o RGB. O melhor resultado obtido para *pixels* corretamente classificados foi de 88,93% no classificador de mistura de Gaussianas, enquanto o melhor resultado para pixels incorretamente classificados foi de 61,90% no classificador de análise de componentes principais.

A avaliação do segundo classificador é feita considerando a sensibilidade, a especificidade e a acurácia, e a área embaixo da curva ROC (*Receiver Operating Characteristics*). O melhor resultado de sensibilidade obtido foi de 94,08% no classificador de k -vizinhos-mais-próximos, a melhor acurácia obtida foi de 93,47% no classificador de mistura de Gaussianas e a melhor especificidade foi 95,85% também no classificador de mistura de Gaussianas. O mapeamento pelo discriminante linear de Fisher gerou resultados piores, de forma geral.

Sotaquirá, Rueda e Narvaez (2009) propõem a utilização da componente Cr do espaço de cores YCbCr e a componente a do espaço de cores Lab. A segmentação é feita em cada uma das componentes separadamente usando uma limiarização global cujo valor é escolhido a partir da primeira derivada do histograma. Não há detalhes sobre os parâmetros utilizados. Em seguida, as imagens segmentadas são combinadas por meio de uma operação AND. O autor argumenta que a componente Cr contém a maior parte da informação dos bacilos e rejeita a maioria dos outros artefatos, mas contém componentes resultantes das condições de variação de luz, enquanto a componente a é mais robusta às variações de iluminação, mas incapaz de rejeitar outros artefatos. Os resultados avaliam a eficiência da segmentação, a sensibilidade e a especificidade do diagnóstico e a acurácia na determinação do nível de infecção. A eficiência da segmentação (definida como o número de bacilos segmentados sobre o número de bacilos totais na imagem) é de 96,3%, enquanto a taxa de falsos positivos (artefatos segmentados sobre o número total de objetos segmentados) é de 9,78%. A sensibilidade e a especificidade do diagnóstico foram de 90,9% e 100%, respectivamente. Essa avaliação foi feita por paciente, avaliando-se 100 imagens de cada paciente, num total de 14 pacientes. Finalmente a avaliação do nível de infecção obteve uma acurácia de 85,7%. O número médio de bacilos foi calculado para cada uma das imagens, e o grau de infecção calculado segundo o método apresentado no artigo.

Makkapati, Agrawal e Acharya (2009) propõem a segmentação das imagens utilizando o método de Otsu seguido de uma análise do número de pixels segmentados na imagem inteira. Caso o número de pixels seja superior a um limite estabelecido, o método de Otsu é aplicado na parte segmentada da imagem. Para a determinação da classe das estruturas segmentadas, buscou-se a presença de uma estrutura nuclear dentro da estrutura segmentada, que seria decorrente da coloração não-uniforme dos bacilos. Finalmente, diferencia-se bacilos de agrupamentos de bacilos baseados nas medidas de comprimento e largura definidas pelo perímetro e a área. Não é apresentado nenhum resultado quantitativo, mas apenas com base na análise visual das imagens.

Nayak, Shenoy e Galigekere (2010) propõem a segmentação supervisionada usando uma classificação de distância mínima entre a amostra e o *cluster* que representa bacilos e entre os *clusters* que representam o fundo. A medida de distância utilizada neste caso é a distância de Mahalanobis, e a análise da segmentação é feita no espaço de cores HSI.

Em seguida, é feita a eliminação de estruturas segmentadas muito pequenas e então faz-se a rotulagem dos elementos conectados, considerando estruturas 8-conectadas. Por fim, estruturas muito próximos são fundidas para diferenciar bacilos de agrupamentos de bacilos. O resultado é medido avaliando a eficiência do método (número de bacilos detectados sobre o número de bacilos totais), e o melhor resultado obtido sobre o conjunto de teste é de 93,48% para bacilos isolados. É importante salientar que a eficiência foi calculada levando em conta o número de bacilos encontrados na imagem, de modo que estruturas que não são bacilos podem ter sido contabilizadas e estruturas que são bacilos podem ter ficado de fora, sem que houvesse alteração do resultado.

Zhai et al. (2010) propõem um método de focalização automática e de reconhecimento e contagem de bacilos em imagens de microscopia convencional. A segmentação é feita usando uma imagem da componente H do espaço de cores HSI utilizando limiarização global definida pelas intensidades de cinza entre 270° e 360° , e a componente L do espaço de cores Lab, onde o limiar é local e definido por uma máscara Gaussiana cujo tamanho é determinado de forma empírica. As duas imagens segmentadas são então combinadas por uma operação lógica AND, seguido de uma operação morfológica de erosão para a eliminação de partículas pequenas. A classificação das estruturas é feita usando uma árvore de decisão onde analisa-se a área das estruturas, a rugosidade (uma razão entre o perímetro do objeto e o perímetro do fecho convexo) e a circularidade, classificando as estruturas em bacilos, agrupamento de bacilos e não-bacilos. A contagem dos bacilos isolados é direta, enquanto a contagem dos bacilos em agrupamentos é feita dividindo a área do agrupamento por um valor médio. A avaliação dos resultados é feita levando em conta a eficiência do método na contagem das estruturas, o que pode ser enganoso, haja vista a existência de falsos positivos que podem mascarar falsos negativos, e medindo a sensibilidade e a especificidade no diagnóstico de cada imagem. Os resultados obtidos foram de 100% e 94%, respectivamente.

Rulaningtyas, Suksmono e Mengko (2011) propõem a classificação das estruturas segmentadas utilizando perímetro, área, comprimentos do maior e menor eixo, circularidade, compacidade, excentricidade e tortuosidade usando uma rede neural artificial do tipo *perceptron* multicamadas com uma camada oculta de 20 neurônios. Os resultados foram avaliados sobre o próprio conjunto de treino com um erro quadrático médio de 0,000368.

Siena et al. (2012) propõem o pré-processamento das imagens de microscopia convencional usando *decorrelation stretch* seguido da segmentação usando *k-means clustering*. A classificação das estruturas é feita utilizando uma rede neural artificial do tipo *perceptron* multicamadas com uma camada escondida de 15 neurônios analisando a excentricidade e a compacidade das estruturas. O resultado alcançado é uma acurácia de 88%. As imagens utilizadas são do CDC - *Centers for Disease Control and Prevention* e consiste em 9 imagens, enquanto para o treinamento da rede, são usadas imagens de bacilos utilizadas

no trabalho de Forero-Vargas et al. (2002).

Costa Filho et al. (2012) propõem a segmentação usando características dos espaços de cores RGB, HSI, YCbCr e Lab em uma rede neural do tipo perceptron multicamadas. Para a classificação das estruturas segmentadas, diversas características foram analisadas e a excentricidade foi escolhida como a característica mais relevante, dentre área, perímetro, compacidade, excentricidade e os três primeiros momentos invariantes. As imagens foram divididas em dois grupos, o primeiro com alta densidade de conteúdo de fundo e o segundo com baixa densidade de conteúdo de fundo. Para imagens de alta densidade de conteúdo de fundo, apenas a excentricidade foi suficiente para produzir bons resultados, mas para as imagens de baixa densidade de conteúdo de fundo, um segundo parâmetro, definido como taxa de cor (CR – *color ratio*) foi utilizado em uma árvore de decisão. Os resultados alcançados considerando todas as imagens foram 91,53%, 91,49% e 8,51% para a sensibilidade, acurácia e falsa detecção, respectivamente.

Kusworo et al. (2013) propõem a segmentação das imagens usando o método de Otsu aplicado à imagem da componente de matiz (H) do espaço de cores HSV, seguido das operações morfológicas de fechamento e abertura. Para a classificação das estruturas, a excentricidade, a compacidade e a métrica (um parâmetro que quantifica quão redondo é uma estrutura) são utilizadas em uma rede neural do tipo *perceptron* multicamadas. A contagem do número de bacilos é feita e comparada com a contagem real feita por um especialista. Neste caso, o resultado medido é o coeficiente de correlação, cujo valor alcançado foi de 0,990. É interessante notar que duas das nove imagens utilizadas foram excluídas da quantificação uma vez que a segmentação proposta não produziu bons resultados. As imagens utilizadas são as mesmas utilizadas por Siena et al. (2012).

Chayadevi e Raju (2014) propõem a segmentação usando o algoritmo de *watershed* nos espaços de cores YCbCr, HSI e Lab. Em seguida, a análise de diversas características é feita para a classificação das estruturas segmentadas. As características analisadas são: área, perímetro, compacidade, excentricidade, eixos maior e menor, a razão entre os eixos, os sete primeiros momentos invariantes e o décimo primeiro momento invariante. Nenhuma classificação e conseqüentemente, nenhuma avaliação quantitativa é realizada.

Embora muitos, a maioria dos trabalhos lida com bancos de imagens diferentes. Ainda, vários trabalhos não avaliam quantitativamente os resultados. Isso faz com que a comparação entre os diversos trabalhos seja difícil. Esse trabalho utiliza o banco de dados usado por Costa et al. (2008) e Costa Filho et al. (2012), o que torna possível uma comparação entre os resultados obtidos. Assim como a maior parte dos trabalhos, um sistema de reconhecimento de bacilos com dois estágios (segmentação da imagem e classificação das estruturas) é proposto. A contribuição desta pesquisa é no sentido de propor uma segmentação eficiente de baixo custo computacional através de uma projeção da imagem baseada na análise de discriminante linear e a classificação das estruturas

usando parâmetros comumente usados na literatura, como os momentos invariantes, a excentricidade e a circularidade, junto de outros associados à cor das estruturas.

1.3 Estrutura do trabalho

O trabalho está dividido da seguinte forma: no capítulo 1 são definidos o problema a ser tratado, a motivação e os objetivos do trabalho, bem como uma revisão da literatura sobre o assunto; no capítulo 2 o referencial teórico utilizado como base para o desenvolvimento do trabalho é apresentado; no capítulo 3 a metodologia adotada na pesquisa é definida; no capítulo 4 os resultados obtidos são apresentados; no capítulo 5 é feita uma análise dos resultados obtidos; e finalmente, no capítulo 6, a conclusão do trabalho é apresentada.

2 Referenciais teóricos

A ideia básica do sistema desenvolvido é apresentada a seguir: primeiramente, a imagem colorida é projetada em uma imagem em intensidade de cinza. A escolha da projeção é obtida por meio da análise do discriminante linear. Em seguida, a segmentação da imagem projetada, agora em níveis de cinza, é feita. Para isso, foram investigadas duas metodologias: uma baseada em limiarização global e outra baseada no método de Otsu. Após a segmentação, um pós-processamento é feito sobre as imagens binárias visando eliminar áreas e melhorar a forma das estruturas segmentadas. Finalmente, a classificação das estruturas é feita. Para tanto, a utilização de redes neurais artificiais e máquinas de vetores suporte considerando vinte características para descrever as estruturas foi analisada. As características que apresentaram os melhores resultados após testes exaustivos foram a excentricidade, a circularidade, sete momentos invariantes, cinco descritores de Fourier e a média e o desvio padrão de cada um dos três canais de cor.

O diagrama de blocos mostrado na Figura 1 representa o sistema de reconhecimento automático de bacilos de tuberculose proposto.



Figura 1 – Diagrama de blocos do sistema desenvolvido.

A seguir, os fundamentos teóricos de cada um dos métodos utilizados são apresentados, bem como a motivação para a escolha de cada um deles.

2.1 Análise de discriminante linear

A análise de discriminante linear, também conhecida como discriminante linear de Fisher, para o caso de duas classes, busca identificar o vetor de projeção \mathbf{v} que maximiza a distância entre as médias das classes (neste caso, *pixels* de bacilo e *pixels* de fundo) e, ao mesmo tempo, minimiza o espalhamento dos dados \mathbf{p} em torno de suas médias, tornando-os mais concentrados ao redor de suas respectivas médias (DUDA; HART; STORK, 2001). Dessa forma, tendo como inspiração a operação de (R – G) feita em Costa et al. (2008) devido à sua capacidade de separação dos *pixels* de bacilo e de fundo e seu baixo custo computacional, a motivação para a utilização de análise de discriminante linear foi encontrar uma projeção da imagem colorida em uma imagem de intensidade de cinza que separasse de forma satisfatória as classes de *pixels* de bacilo e *pixels* de fundo.

Esse objetivo é alcançado maximizando a seguinte função critério, também conhecida como razão de discriminante de Fisher:

$$J = \frac{|m_1(\mathbf{v}) - m_2(\mathbf{v})|^2}{s_1(\mathbf{v})^2 + s_2(\mathbf{v})^2}, \quad (2.1)$$

onde $m_1(\mathbf{v})$ e $m_2(\mathbf{v})$ representam as médias da projeção das duas classes e $s_1(\mathbf{v})$ e $s_2(\mathbf{v})$ representam os espalhamentos em torno das médias dos dados projetados das duas classes.

Considerando \mathbf{v} o vetor de projeção dos dados descritos pelo vetor das observações \mathbf{x} , a projeção é então dada por

$$u = \mathbf{v}^t \mathbf{p}. \quad (2.2)$$

Para o caso de duas classes χ_1 e χ_2 , com médias $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$, respectivamente, as médias da projeção das duas classes são descritas como

$$\begin{aligned} m_i(\mathbf{v}) &= \frac{1}{n_i} \sum_{\mathbf{p} \in \chi_i} u \\ &= \frac{1}{n_i} \sum_{\mathbf{p} \in \chi_i} \mathbf{v}^t \mathbf{p} \\ &= \mathbf{v}^t \boldsymbol{\mu}_i. \end{aligned} \quad (2.3)$$

Dessa maneira, para \mathbf{p} pertencente a \mathbb{R}^3 , como é o caso de cada *pixel*,

$$\begin{aligned} |m_1(\mathbf{v}) - m_2(\mathbf{v})|^2 &= (m_1(\mathbf{v}) - m_2(\mathbf{v}))^2 \\ &= (\mathbf{v}^t \boldsymbol{\mu}_1 - \mathbf{v}^t \boldsymbol{\mu}_2)^2 \\ &= \mathbf{v}^t (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^t \mathbf{v} \\ &= \mathbf{v}^t \mathbf{S}_B \mathbf{v}, \end{aligned} \quad (2.4)$$

e

$$\mathbf{S}_B = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^t. \quad (2.5)$$

\mathbf{S}_B é definida como a matriz de espalhamento entre-classes, pois representa uma medida da distância entre os centros das duas classes.

De forma semelhante,

$$\begin{aligned}
 s_i^2 &= \sum_{u \in \chi_i} (u - m_i)^2 \\
 &= \sum_{\mathbf{p} \in \chi_i} (\mathbf{v}^t \mathbf{p} - \mathbf{v}^t \boldsymbol{\mu}_i)^2 \\
 &= \sum_{\mathbf{p} \in \chi_i} \mathbf{v}^t (\mathbf{p} - \boldsymbol{\mu}_i) (\mathbf{p} - \boldsymbol{\mu}_i)^t \mathbf{v} \\
 &= \mathbf{v}^t \mathbf{S}_i \mathbf{v},
 \end{aligned} \tag{2.6}$$

onde

$$\mathbf{S}_i = \sum_{\mathbf{p} \in \chi_i} (\mathbf{p} - \boldsymbol{\mu}_i) (\mathbf{p} - \boldsymbol{\mu}_i)^t. \tag{2.7}$$

Dessa forma,

$$\begin{aligned}
 s_1^2 + s_2^2 &= \mathbf{v}^t \mathbf{S}_1 \mathbf{v} + \mathbf{v}^t \mathbf{S}_2 \mathbf{v} \\
 &= \mathbf{v}^t (\mathbf{S}_1 + \mathbf{S}_2) \mathbf{v} \\
 &= \mathbf{v}^t \mathbf{S}_W \mathbf{v}.
 \end{aligned} \tag{2.8}$$

$\mathbf{S}_W = \mathbf{S}_1 + \mathbf{S}_2$ é definida como a matriz de espalhamento intra-classes.

Com as matrizes de espalhamento entre-classes e intra-classes, é possível escrever a função critério em relação ao vetor de projeção \mathbf{v} :

$$J(\mathbf{v}) = \frac{\mathbf{v}^t \mathbf{S}_B \mathbf{v}}{\mathbf{v}^t \mathbf{S}_W \mathbf{v}}. \tag{2.9}$$

Essa função critério é conhecida como o quociente de Rayleigh generalizado, cujo vetor \mathbf{v} que maximiza $J(\mathbf{v})$ é a solução de um problema de autovalores generalizado:

$$\mathbf{S}_B \mathbf{v} = \lambda \mathbf{S}_W \mathbf{v}. \tag{2.10}$$

Uma solução que maximiza $J(\mathbf{v})$ é

$$\mathbf{v} = \mathbf{S}_W^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2), \tag{2.11}$$

de forma que o cálculo da projeção \mathbf{v} depende da inversão da matriz \mathbf{S}_W .

O propósito da utilização de análise de discriminante linear é reduzir a dimensionalidade do problema de tal forma que as duas classes possuam uma separação maior no sentido do critério de Fisher, $J(\mathbf{v})$, ou seja, uma maior distância entre as suas médias e uma menor dispersão dos dados de cada classe em torno de suas respectivas médias. No caso deste trabalho, a redução da dimensionalidade proporciona a representação de uma imagem colorida em uma imagem em níveis de cinza.

Mais detalhes sobre a análise de discriminante linear podem ser encontrados em Duda, Hart e Stork (2001) e Theodoridis e Koutroumbas (2009).

2.2 Segmentação das imagens

A segmentação de imagens é o processo de dividir uma imagem em regiões ou objetos que a compõem (GONZALEZ; WOODS, 2007). Embora várias classes possam ser obtidas do processo de segmentação, apenas a segmentação que divide a imagem em duas classes será discutida por se aplicar ao problema em questão. Além disso, apenas a segmentação de imagens de intensidade de cinza será tratada, uma vez que a projeção feita com base na análise do discriminante linear transforma as imagens coloridas em imagens de intensidade de cinza.

2.2.1 Segmentação por limiarização global

A segmentação por limiarização global divide a imagem em duas classes baseando-se simplesmente na intensidade de cinza de cada *pixel*, $f(x, y)$, conforme a seguinte regra (GONZALEZ; WOODS, 2007):

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) \geq T \\ 0 & \text{se } f(x, y) < T \end{cases}, \quad (2.12)$$

onde $g(x, y)$ é o valor do *pixel* na imagem segmentada e T é o valor de limiar que separa as duas classes.

Dessa forma, todos os *pixels* cuja intensidade de cinza seja maior ou igual a T são classificadas em uma classe e os *pixels* cuja intensidade de cinza é menor que T são classificados em outra.

2.2.2 Método de Otsu

O método de Otsu é um procedimento para a segmentação de imagens de intensidade de cinza que busca encontrar um limiar T que maximiza a separabilidade entre as classes resultantes (OTSU, 1979).

O método consiste em implementar a técnica de limiarização global para todos os valores possíveis de intensidade de cinza (por exemplo, de 0 a 255 no caso de imagens representadas em 8 bits) e avaliar, para as duas classes geradas, a função critério

$$\sigma^2 = P_1 P_2 (\nu_1 - \nu_2)^2, \quad (2.13)$$

onde P_1 e P_2 são as probabilidades de que um *pixel* seja atribuído às classes 1 ou 2, respectivamente, e ν_1 e ν_2 representam a média da intensidade de cinza de cada classe. Esses valores são calculados como

$$P_1(T) = \sum_{i=0}^T q_i, \quad (2.14)$$

$$P_2(T) = 1 - P_1(T), \quad (2.15)$$

$$\nu_1(T) = \frac{1}{P_1(T)} \sum_{i=0}^T i q_i, \quad (2.16)$$

$$\nu_2(T) = \frac{1}{P_2(T)} \sum_{i=T+1}^{L-1} i q_i, \quad (2.17)$$

onde L é o número de possíveis valores de intensidade de cinza e q_i é o número de *pixels* que possuem a intensidade de cinza i dividido pelo número total de *pixels* da imagem. O limiar T escolhido é aquele que maximiza a função critério σ^2 .

2.3 Operações morfológicas em imagens digitais binárias

Operadores morfológicos binários são de natureza não linear, necessitam da definição de um elemento estruturante, e tem como operações básicas a dilatação e a erosão binária. A opção por usar um operador não linear, porém algébrico, permite que tanto a imagem de entrada como a imagem de saída sejam binárias. Dessa forma, para eliminar ruídos com formato de pequenos vales ou pequenas protuberâncias após a etapa de segmentação, operações morfológicas foram utilizadas sobre as imagens segmentadas.

As operações morfológicas utilizadas neste trabalho são descritas a seguir.

2.3.1 Operação de abertura

A abertura de uma imagem I pelo elemento estruturante E é definida como uma operação de erosão de I por E seguida de uma dilatação do resultado por E (GONZALEZ; WOODS, 2007), sendo representada por

$$I \circ E = (I \ominus E) \oplus E, \quad (2.18)$$

onde \circ , \ominus , e \oplus representam as operações de abertura, erosão e dilatação, respectivamente.

De forma geral, a abertura de uma imagem suaviza o contorno dos objetos, rompe os istmos (pequenas protuberâncias) e elimina pequenas saliências (GONZALEZ; WOODS, 2007).

2.3.2 Operação de fechamento

O fechamento de uma imagem I pelo elemento estruturante E é definido como uma operação de dilatação de I por E seguida de uma erosão do resultado por E (GONZALEZ; WOODS, 2007), sendo representado por

$$I \bullet E = (I \oplus E) \ominus E, \quad (2.19)$$

onde \bullet representa a operação de fechamento.

Assim como a operação de abertura, a operação de fechamento geralmente suaviza o contorno dos objetos, mas funde as descontinuidades estreitas e elimina pequenos buracos preenchendo as lacunas dentro de um contorno (GONZALEZ; WOODS, 2007).

2.4 Representação de estruturas em imagens digitais

O objetivo aqui é representar cada bacilo segmentado por um vetor que o caracterize, sendo esse vetor de dimensão menor que a quantidade de *pixels* que compõe o bacilo. Obtém-se assim tanto uma redução dimensional como uma caracterização robusta dos dados, uma vez que formas diversas de bacilos serão representadas por pontos no espaço n -dimensional. Por meio de teste sucessivos, optou-se por utilizar as vinte características descritas a seguir por, em conjunto, produzirem os melhores resultados.

2.4.1 Excentricidade

A excentricidade é uma medida diretamente relacionada com o alongamento das estruturas tendo sido utilizados nos trabalhos de Sadaphal et al. (2008), Khutlang et al. (2010), Rulaningtyas, Suksmono e Mengko (2011), Siena et al. (2012), Costa Filho et al. (2012), Kusworo et al. (2013) e Chayadevi e Raju (2014).

A excentricidade é um valor que especifica a excentricidade de uma elipse que possui o mesmo segundo momento da estrutura. No caso da elipse, a excentricidade é a razão entre o foco da elipse e o comprimento do maior eixo, de modo que um círculo possui excentricidade 0 e um segmento de linha tem excentricidade igual a 1.

A Figura 2 mostra uma elipse no plano cartesiano onde estão definidos o eixo maior, a , sobre o eixo das abscissas, e o eixo menor, b , sobre o eixo das ordenadas. A excentricidade é calculada fazendo

$$Ex = \sqrt{1 - \frac{b^2}{a^2}}. \quad (2.20)$$

2.4.2 Circularidade

A circularidade, ou razão de circularidade, é uma medida da semelhança entre a estrutura e um círculo e é definida como

$$C = \frac{4\pi A}{Q^2}, \quad (2.21)$$

e pode ser entendida como uma razão entre a área, A , de uma região e a área de um círculo com o mesmo perímetro, Q , da região (GONZALEZ; WOODS, 2007). Dessa maneira, para

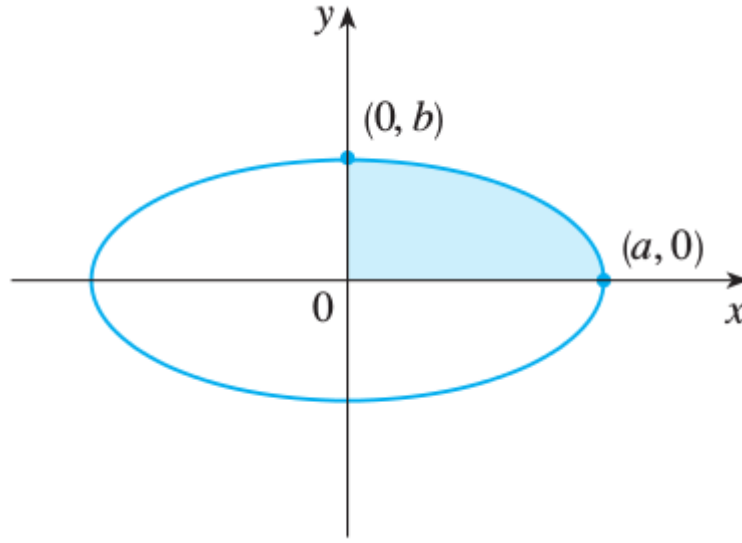


Figura 2 – Elipse, com destaque para os eixos maior e menor. Fonte: Stewart (2012)

um círculo, a circularidade se iguala a 1. A circularidade foi utilizada para a classificação das estruturas em Forero-Vargas et al. (2002), Khutlang et al. (2010), Zhai et al. (2010), Rulaningtyas, Suksmono e Mengko (2011), Siena et al. (2012), Costa Filho et al. (2012), Kusworo et al. (2013) e Chayadevi e Raju (2014).

2.4.3 Momentos invariantes

Os momentos invariantes de uma estrutura, ou de uma região, representam uma medida da distribuição dos *pixels* em torno de seu centro de massa. Eles são invariantes às operações de mudança de escala, rotação e translação (GONZALEZ; WOODS, 2007) e estão entre as mais importantes medidas de descrição de estruturas, uma vez que as formas de estruturas no espaço de duas dimensões estão diretamente relacionadas com seus momentos (LEU, 1991).

Na classificação de bacilos de tuberculose, os trabalhos de Khutlang et al. (2010), Costa Filho et al. (2012) e Chayadevi e Raju (2014) mostraram que momentos invariantes são bons descritores de forma das estruturas segmentadas.

O momento 2D de ordem $(p + q)$ de uma estrutura cuja intensidade dos *pixels* é dada por $f(x, y)$ é:

$$\rho_{pq} = \sum_x \sum_y x^p y^q f(x, y). \quad (2.22)$$

O centro de massa pode então ser localizado por

$$\bar{x} = \frac{\rho_{10}}{\rho_{00}} \text{ e } \bar{y} = \frac{\rho_{01}}{\rho_{00}}. \quad (2.23)$$

Os momentos centrais de ordem $(p + q)$ são definidos por

$$\kappa_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y). \quad (2.24)$$

Os momentos centrais normalizados em relação à intensidade dos *pixels* da região ou da estrutura em questão são calculados como

$$\eta_{pq} = \frac{\kappa_{pq}}{\kappa_{00}^v}, \quad (2.25)$$

onde

$$v = \frac{p + q}{2} + 1. \quad (2.26)$$

Em Hu (1962) são definidos sete momentos invariantes independentes da posição, escala e orientação das estruturas:

$$\phi_1 = \eta_{20} + \eta_{02}, \quad (2.27)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \quad (2.28)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \quad (2.29)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2, \quad (2.30)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \end{aligned} \quad (2.31)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \quad (2.32)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned} \quad (2.33)$$

2.4.4 Descritores de Fourier

Descritores de Fourier são descritores de forma baseados na fronteira de uma estrutura e representam uma medida indireta desta. Na classificação de bacilos de tuberculose, os descritores de Fourier foram utilizados no reconhecimento de bacilos de tuberculose em imagens de microscopia por fluorescência em Veropoulos et al. (1998) e Forero-Vargas et al. (2002), e em imagens de microscopia convencional em Khutlang et al. (2010).

Os pontos da fronteira (x, y) são mapeados na forma de um número complexo

$$r(k) = x(k) + jy(k), \quad (2.34)$$

onde $k = 0, 1, 2, \dots, (K - 1)$ são os pontos da fronteira, K é o número total de *pixels* que constituem a fronteira e $j = \sqrt{-1}$.

A transformada discreta de Fourier de $r(k)$ é então dada por

$$c(o) = \sum_{k=0}^{K-1} r(k) e^{\frac{-j2\pi ok}{K}}, \quad (2.35)$$

onde o é a variável do domínio da frequência.

O módulo dos coeficientes complexos $c(o)$ são então os descritores de Fourier da fronteira em questão. Normalmente, apenas alguns poucos descritores de Fourier são necessários para representar de forma adequada as principais características de uma fronteira (GONZALEZ; WOODS, 2007). Os descritores de Fourier apresentados na Equação (2.35) não são invariantes à rotação, translação e escala. Uma forma de torná-los invariantes a essas operações é trabalhar com os seguintes descritores, conforme indicado por Costa e Cesar (2000):

$$F(o) = |N_F(o)|, \quad (2.36)$$

onde

$$N_F(o) = \begin{cases} 0 & \text{se } o = 0 \\ \frac{c(o)}{c(1)} & \text{se } o \neq 0 \end{cases}. \quad (2.37)$$

2.4.5 Descritores de cores

Descritores de cores são medidas matemáticas que buscam descrever objetos ou estruturas conforme seus atributos de cor. Neste trabalho, apenas a média e o desvio padrão de cada canal de cor no espaço de cores RGB são calculados para cada estrutura.

A motivação para sua utilização está na ideia de que a média dos *pixels* de cada estrutura, bem como a distribuição dos *pixels* em torno desta para cada um dos canais de cor, contém informação relevante a ser usada pelos classificadores.

Constatou-se que a utilização dessas seis características relacionadas à cor eleva o desempenho dos classificadores em comparação à classificação sem considerá-las.

2.5 Classificadores

Uma vez que se encontra associado aos bacilos um vetor de característica que os represente, é necessário rotular cada um desses vetores como bacilos ou não bacilos. Para isso, a seguir, serão discutidos classificadores de padrões, supervisionados, que apresentaram bons resultados para o problema em questão.

2.5.1 Redes neurais artificiais

Em Haykin et al. (2009), na página 2, a seguinte definição para redes neurais artificiais é apresentada (tradução livre):

"Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso".

Elas recebem seu nome devido à sua inspiração no processamento realizado pelo cérebro. As redes neurais possuem diversas propriedades úteis tais como a não-linearidade e a adaptabilidade, o que justifica sua grande aplicação em problemas de reconhecimento de padrões.

A unidade básica de processamento de uma rede neural é o neurônio que, a partir de um vetor de entrada \mathbf{x} gera a saída y_k através da operação descrita por

$$y_k = \varphi(\mathbf{w}_k^t \mathbf{x} + b_k), \quad (2.38)$$

onde \mathbf{w}_k é o vetor dos pesos relativos a cada entrada do neurônio; b_k é um *offset*, cujo efeito é aumentar ou diminuir o escalar obtido por $\mathbf{w}_k^t \mathbf{x}$; e $\varphi(\cdot)$ é a função de ativação que limita o intervalo permissível de amplitude do sinal de saída a um valor finito. A Figura 3 mostra o modelo de um neurônio que realiza a operação descrita pela Equação (2.38).

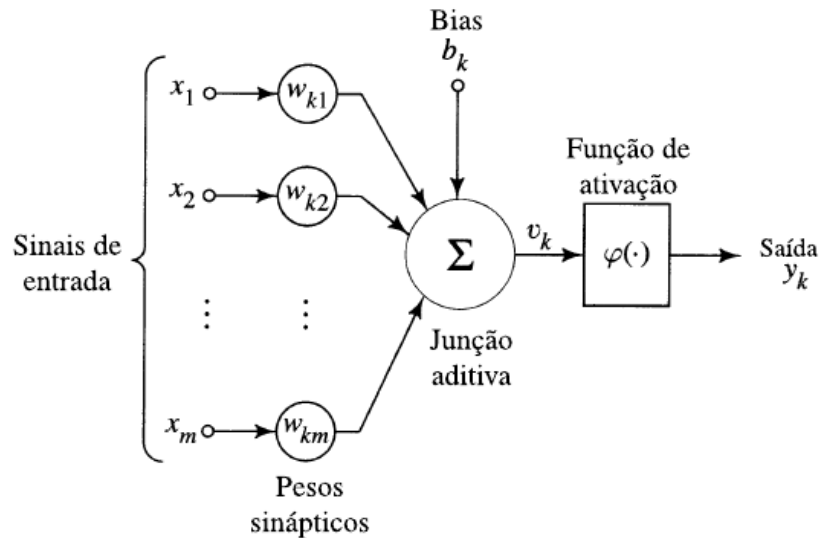


Figura 3 – Modelo de um neurônio. Fonte: (HAYKIN et al., 2009)

Em alguns casos, também é comum representar o neurônio conforme o modelo apresentado na Figura 4 onde o *offset* é representado como uma entrada igual a 1 associada a um peso w_{k0} . Nesta representação a operação realizada por cada unidade de processamento

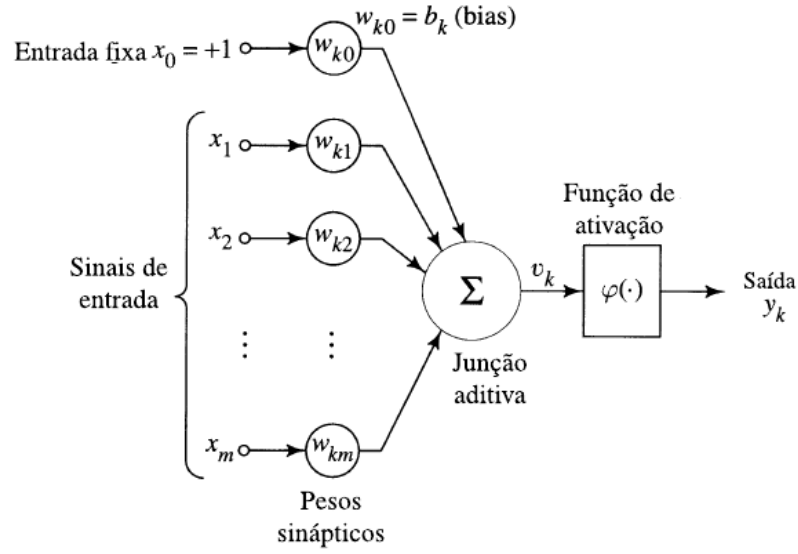


Figura 4 – Modelo alternativo de um neurônio. Fonte: (HAYKIN et al., 2009)

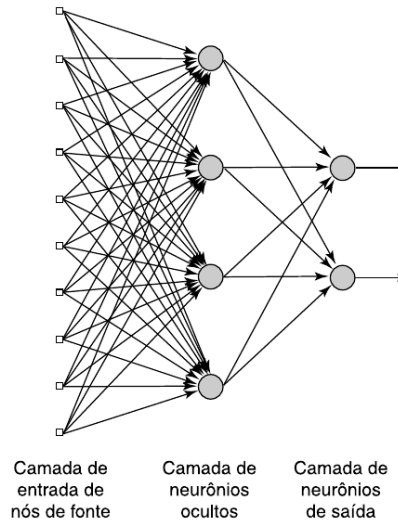


Figura 5 – Rede neural alimentada diretamente com uma única camada oculta. Fonte: (HAYKIN et al., 2009)

é simplesmente

$$y_k = \varphi(\mathbf{w}_k^t \mathbf{x}), \quad (2.39)$$

onde o *offset* é inserido em \mathbf{w} e associado à entrada -1 , também inserida em \mathbf{x} .

As redes neurais possuem diversas arquiteturas, sendo que a rede neural alimentada diretamente com múltiplas camadas é uma das arquiteturas mais comuns. Uma rede alimentada diretamente com uma única camada oculta é mostrada na Figura 5, onde cada círculo representa um neurônio cujo modelo é mostrado na Figura 3. A rede neural mostrada na Figura 5 possui 10 entradas, ou \mathbf{x} é um vetor com 10 dimensões, 4 neurônios na camada oculta, e 2 neurônios na camada de saída. Esse modelo também é conhecido como *perceptron* de múltiplas camadas.

A saída de cada camada da rede neural mostrada na Figura 5 é dada pela operação matricial

$$\mathbf{y} = \varphi(\mathbf{W}^t \mathbf{x}), \quad (2.40)$$

onde \mathbf{y} é o vetor coluna que contém a saída de cada um dos neurônios da camada em questão, \mathbf{x} é o vetor de entrada dessa camada, e \mathbf{W} é a matriz que contém o peso de cada neurônio relativo a cada entrada. Neste caso, cada coluna de \mathbf{W} é um vetor \mathbf{w}_k como mostrado na Equação (2.39). Admite-se que tanto a entrada quanto os pesos dos neurônios estão na forma do modelo da Figura 4, e $\varphi(\cdot)$ é realizada sobre cada termo do vetor que resulta da operação $(\mathbf{W}^t \mathbf{x})$.

Redes neurais artificiais da forma mostrada na Figura 5 podem ser usadas como aproximadores universais de funções, ou ainda, como classificadores de padrões. O treinamento de uma rede neural do tipo *perceptron* de múltiplas camadas é supervisionado, isto é, o treinamento é feito com base em dados de entrada juntamente com as saídas reais correspondentes, e, geralmente, envolve a minimização do erro quadrático médio entre os valores obtidos pela rede e os valores reais relativos a cada padrão de entrada.

Nesse contexto, tem destaque o algoritmo de Levenberg-Marquardt para o treinamento por retropropagação, devido a sua alta eficiência no treinamento de redes com até algumas centenas de pesos (HAGAN; MENHAJ, 1994).

2.5.2 Máquinas de vetores suporte

Máquinas de vetores suporte (SVM, do inglês, *support vector machines*) são, em sua essência, classificadores de duas classes. As máquinas de vetores suporte também são conhecidas como máquinas de *kernel* e são métodos de maximização da margem que permitem ao modelo do classificador ser escrito como uma soma das influências de um subconjunto dos dados de treinamento (ALPAYDIN, 2004).

No contexto da classificação das estruturas como bacilos ou não-bacilos, as máquinas de vetores suporte foram utilizadas uma vez que sua principal aplicação é no reconhecimento de duas classes. Além disso, a aplicação de máquinas de vetores suporte é, em geral, mais rápida do que a aplicação de redes neurais artificiais no contexto de classificação, de modo que, no caso de resultados similares, as primeiras podem ser utilizadas visando menor esforço computacional.

Para entender o funcionamento das máquinas de vetores suporte, considere o caso de duas classes χ_1 e χ_2 linearmente separáveis. Cada amostra \mathbf{x} possui um rótulo h tal que

$$h = \begin{cases} +1 & \text{se } \mathbf{x} \in \chi_1 \\ -1 & \text{se } \mathbf{x} \in \chi_2 \end{cases}. \quad (2.41)$$

Neste caso, o objetivo é encontrar $\boldsymbol{\omega}$ e ω_0 , tal que

$$\begin{aligned} \boldsymbol{\omega}^t \mathbf{x} + \omega_0 &\geq +1 & \text{se} & \quad h = +1 \\ \boldsymbol{\omega}^t \mathbf{x} + \omega_0 &\leq -1 & \text{se} & \quad h = -1 \end{aligned} \quad (2.42)$$

A distância de uma amostra \mathbf{x} ao hiperplano do discriminante é

$$\frac{|\boldsymbol{\omega}^t \mathbf{x} + \omega_0|}{\|\boldsymbol{\omega}\|} = \frac{h (\boldsymbol{\omega}^t \mathbf{x} + \omega_0)}{\|\boldsymbol{\omega}\|}. \quad (2.43)$$

É desejável que essa distância seja pelo menos um valor d , de modo a colocar o hiperplano do discriminante em uma posição que permita uma melhor generalização, ou seja, que dados que não pertencem ao conjunto de treinamento possam continuar afastados do discriminante para que não haja dúvidas sobre sua classe (ALPAYDIN, 2004). Nesse caso:

$$\frac{|\boldsymbol{\omega}^t \mathbf{x} + \omega_0|}{\|\boldsymbol{\omega}\|} = \frac{h (\boldsymbol{\omega}^t \mathbf{x} + \omega_0)}{\|\boldsymbol{\omega}\|} \geq d, \forall \mathbf{x}. \quad (2.44)$$

Define-se a margem de separação entre as classes como a distância do hiperplano de separação até as amostras mais próximas a ela de ambos os lados. Dessa forma, maximizar a margem de separação entre as classes equivale a maximizar d . Entretanto, um número infinito de soluções pode aparecer apenas mudando o valor absoluto de $\boldsymbol{\omega}$. Nesse caso, faz-se $h \|\boldsymbol{\omega}\| = 1$, de modo que maximizar d é equivalente a minimizar $\|\boldsymbol{\omega}\|$, ou ainda, minimizar $(1/2)\|\boldsymbol{\omega}\|^2$.

Encontrar o hiperplano de separação equivale então a minimizar

$$\frac{1}{2}\|\boldsymbol{\omega}\|^2, \quad (2.45)$$

sujeita à condição

$$h (\boldsymbol{\omega}^t \mathbf{x} + \omega_0) \geq +1. \quad (2.46)$$

Esse é um problema de otimização quadrática, cuja solução é $\boldsymbol{\omega}$ e ω_0 definidos como uma combinação linear das amostras que estão mais perto do discriminante, os vetores suporte, como ilustra a Figura 6.

A classificação é então feita da seguinte forma

$$\begin{aligned} \mathbf{x} \in \chi_1 & \quad \text{se} \quad \boldsymbol{\omega}^t \mathbf{x} + \omega_0 \geq 0 \\ \mathbf{x} \in \chi_2 & \quad \text{se} \quad \boldsymbol{\omega}^t \mathbf{x} + \omega_0 < 0 \end{aligned} \quad (2.47)$$

No caso mais geral, em que as classes não são linearmente separáveis, a utilização de um discriminante linear tem como consequência a classificação errada de algumas amostras de treinamento. Nesse caso, entende-se por erro não apenas a classificação errônea, mas também quando a distância de \mathbf{x} ao discriminante não é suficientemente grande.

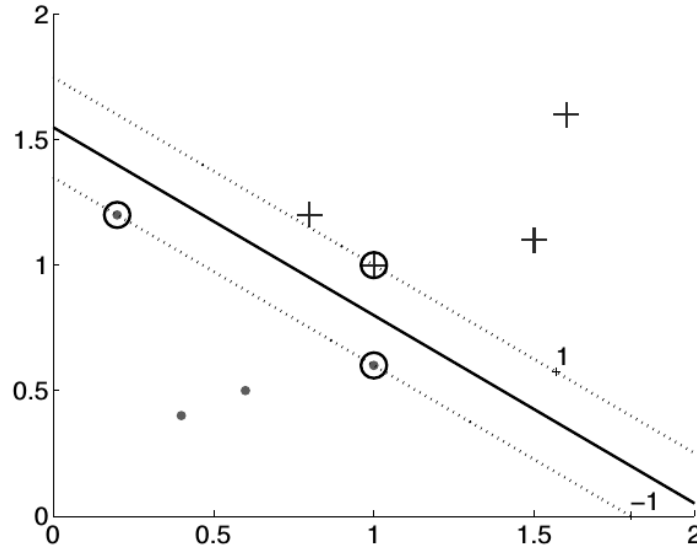


Figura 6 – Vetores suporte em um problema de classificação envolvendo duas classes linearmente separáveis. Fonte: (ALPAYDIN, 2004)

Assim a função a ser minimizada é

$$\frac{1}{2} \|\omega\|^2 + B \sum \xi, \quad (2.48)$$

sujeita às condições

$$h(\omega^t \mathbf{x} + \omega_0) \geq 1 - \xi \quad \text{e} \quad \sum \xi \geq 0, \quad (2.49)$$

onde B é um fator de penalização e ξ é uma variável que quantifica o erro. $\xi = 0$ se \mathbf{x} é corretamente classificado e está suficientemente distante do discriminante; $0 < \xi < 1$ se \mathbf{x} é corretamente classificado, mas não está suficientemente distante do discriminante; e $\xi \geq 1$ se \mathbf{x} é classificado de forma incorreta.

Quando os dados não são linearmente separáveis, uma alternativa é projetar os dados em um novo espaço de maior dimensão onde espera-se que um modelo linear de discriminante possa ser utilizado para separar os dados (ALPAYDIN, 2004). A projeção é dada por

$$\mathbf{z} = \gamma(\mathbf{x}), \quad (2.50)$$

e o discriminante é descrito por

$$\omega^t \mathbf{z} = \omega^t \gamma(\mathbf{x}) \quad (2.51)$$

Como não existem garantias de que as classes serão linearmente separáveis no novo espaço, a função que deve ser minimizada é

$$\frac{1}{2} \|\omega\|^2 + B \sum \xi, \quad (2.52)$$

sujeita às condições

$$h(\omega^t \gamma(\mathbf{x}) + \omega_0) \geq 1 - \xi \quad \text{e} \quad \sum \xi \geq 0, \quad (2.53)$$

Na utilização do método de multiplicadores de Lagrange para resolver o problema de otimização quadrática, a projeção $\gamma(\mathbf{x})$ surge através do termo $\gamma(\mathbf{x})^t \gamma(\mathbf{x})$. A ideia das máquinas de *kernel* é substituir esse produto interno por uma função de *kernel* $\Upsilon(\mathbf{x}, \mathbf{x}')$ que representa uma matriz quadrada de tamanho igual ao conjunto de treinamento.

Algumas funções de *kernel* comumente utilizadas são:

- polinômios de grau n

$$\Upsilon(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^t \mathbf{x}' + 1)^n \quad (2.54)$$

- funções de base radial de raio r

$$\Upsilon(\mathbf{x}, \mathbf{x}') = e^{-\frac{D(\mathbf{x}, \mathbf{x}')}{2r^2}} \quad (2.55)$$

onde $D(\mathbf{x}, \mathbf{x}')$ é uma medida de distância que pode ser, por exemplo, a distância euclidiana ou a distância de Mahalanobis.

- funções sigmoidais

$$\Upsilon(\mathbf{x}, \mathbf{x}') = \tanh(2\mathbf{x}^t \mathbf{x}' + 1) \quad (2.56)$$

3 Metodologia

3.1 Imagens utilizadas no trabalho

As imagens utilizadas neste trabalho são as mesmas usadas em Costa et al. (2008) e Costa Filho et al. (2012) e estão disponíveis em <<http://www.tbimages.ufam.edu.br/>> (CETELI, 2014).

O banco de dados é composto de 120 imagens de 12 pacientes, sendo 10 imagens de uma mesma lâmina para cada paciente. Em todas as imagens, os bacilos assinalados foram identificados por um especialista que marcou os bacilos com círculos; agrupamentos de bacilos foram marcados com retângulos; e estruturas duvidosas que podem ser bacilos, mas das quais não é possível ter certeza devido à incorreta focalização da câmera de captura ou à geometria das estruturas dos supostos bacilos, foram marcadas com polígonos.

As imagens de lâminas de esfregaço de escarro foram preparadas segundo a técnica de Ziehl-Neelsen. As imagens foram obtidas usando uma câmera Canon PowerShot A640 de 10 *MegaPixels* acoplada a um microscópio cujo modelo é Zeiss Axioskop 4.

As imagens estão em formato *JPEG*¹, possuem tamanho de 2816 x 2112 *pixels* e não sofreram nenhum tipo de pré-processamento além da escolha do foco. A Figura 8 mostra algumas das imagens pertencentes ao banco de dados, onde é possível notar a grande diferença entre o fundo e a quantidade de bacilos das imagens. A Figura 9 mostra as mesmas imagens da Figura 8 onde os bacilos foram marcados pelo especialista.

É possível observar nas imagens da Figura 8 a grande variabilidade das imagens no que diz respeito à densidade de estruturas de fundo e a densidade de bacilos. Todas as imagens possuem pelo menos um bacilo, embora muitas imagens contêm poucos bacilos. O histograma apresentado na Figura 7 mostra que mais de 70% das imagens possuem menos de 40 bacilos. Nas Figuras 8 e 7, (a), (b) e (c) representam com grande densidade de estruturas de fundo, enquanto (e), (f) e (g) são imagens com baixa densidade de estruturas de fundo. Todas as imagens, no entanto, possuem grande quantidade de bacilos, segundo a classificação das imagens do banco de dados (CETELI, 2014).

¹ Geralmente, adota-se formatos de compressão sem perdas em imagens para utilizações em medicina. O JPEG é um formato de compressão com perdas, entretanto, os artefatos de compressão não impediram o desenvolvimento de um sistema de reconhecimento automático de bacilos eficiente.

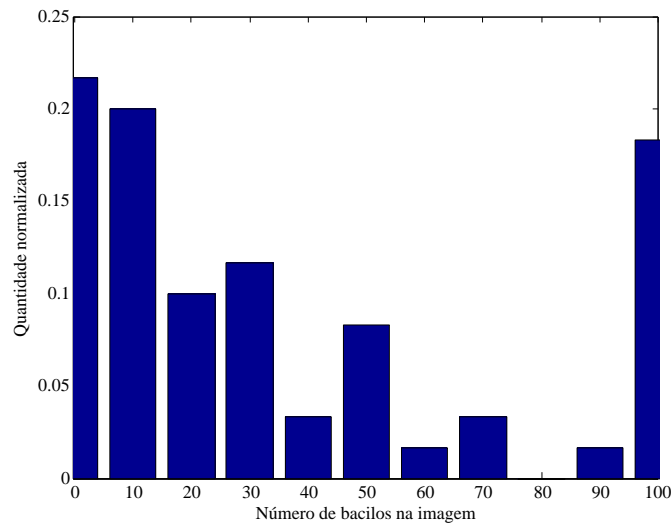


Figura 7 – Histograma das áreas de bacilos.

3.2 Sistema proposto

O método proposto é dividido em dois estágios, a segmentação das imagens e a classificação das estruturas segmentadas.

3.2.1 Segmentação das imagens

Com o objetivo de segmentar os bacilos, cada imagem colorida é projetada em uma imagem em intensidade de cinza baseada na análise do discriminante linear de Fisher, de modo a permitir a utilização de técnicas de segmentação de imagens em níveis de cinza.

Cada *pixel* da imagem colorida pode ser visto como um vetor de três dimensões, onde cada dimensão representa uma componente de cor do espaço de cores RGB. A projeção é então feita da forma $u = \mathbf{v}^t \mathbf{p}$ de modo a obter um único valor para cada *pixel*. O vetor de projeção é obtido pela análise do discriminante linear de Fisher, na qual uma classe representa os *pixels* referentes a bacilos e *pixels* que não pertencem a bacilos.

Nesse caso, a projeção baseada na análise do discriminante de Fisher visa maximizar a separação representada pela Equação (2.1) entre os *pixels* de bacilos e os *pixels* que não são bacilos, gerando uma imagem em intensidade de cinza onde os *pixels* de bacilos estão mais afastados dos *pixels* que não são bacilos.

A projeção foi escolhida de modo a tornar os *pixels* de bacilos mais claros do que os *pixels* das estruturas que não pertencem a bacilos. Em uma imagem normalizada, cujos valores de intensidade se encontram no intervalo $[0, 1]$, os *pixels* de bacilo estão mais próximos de 1 e os *pixels* de fundo estão mais próximos de 0.

A matiz é uma característica associada à cromaticidade de uma cor, ou seja, ao

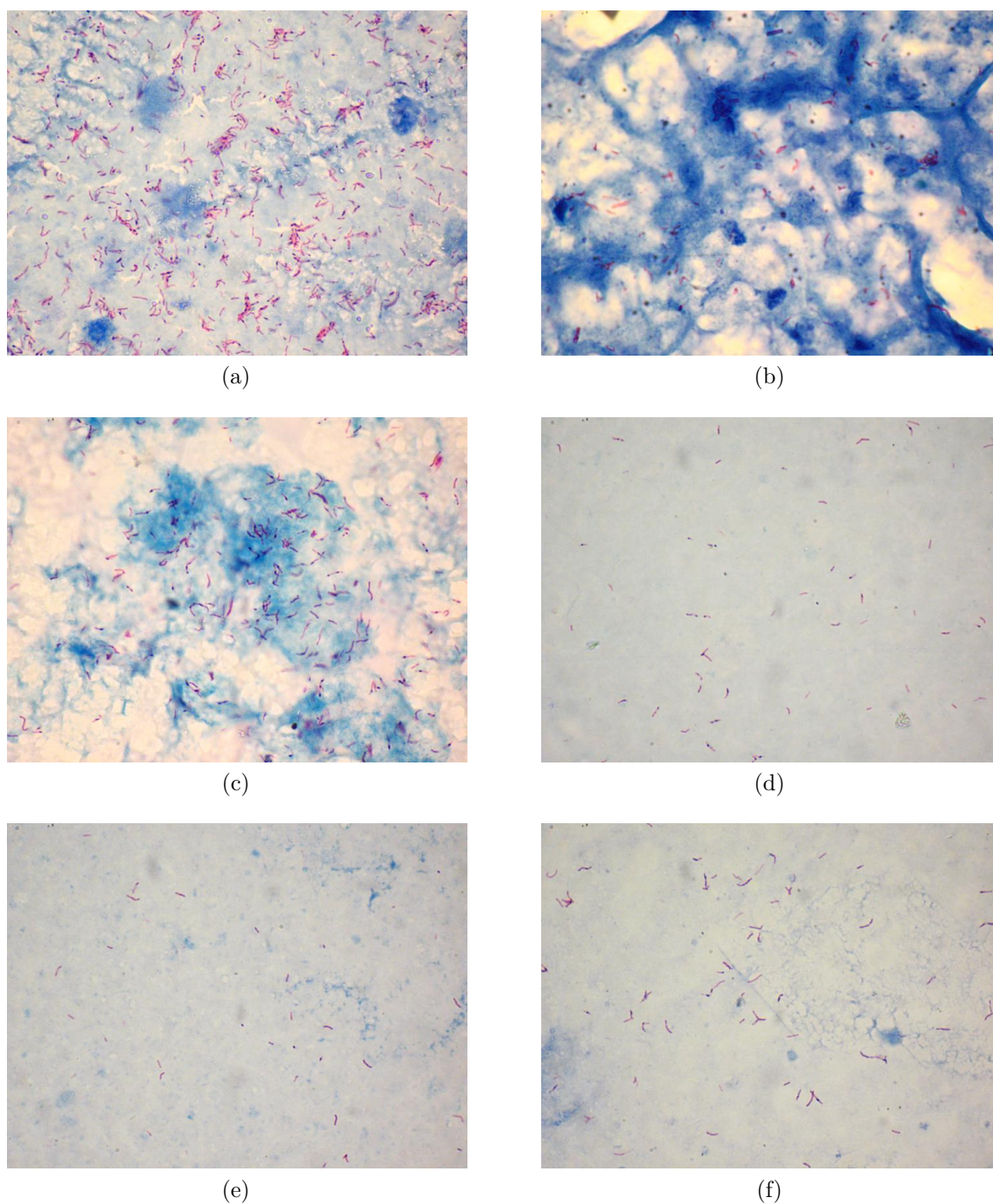


Figura 8 – Exemplos de imagens do banco de dados.

comprimento de onda dominante em uma mistura de ondas de luz e, assim, representa a cor dominante percebida por um observador (KOSCHAN; ABIDI, 2008; GONZALEZ; WOODS, 2007)

Para a análise do discriminante linear de Fisher foram considerados apenas os *pixels* de fundo cujo valor da componente de matiz estava compreendido no intervalo $f(x, y) \leq 0,2$ ou $f(x, y) \geq 0,7$. Os *pixels* pertencentes ao intervalo de matiz $0,2 < f(x, y) < 0,7$ (fora do intervalo utilizado) definitivamente não representam bacilos e, portanto, podem

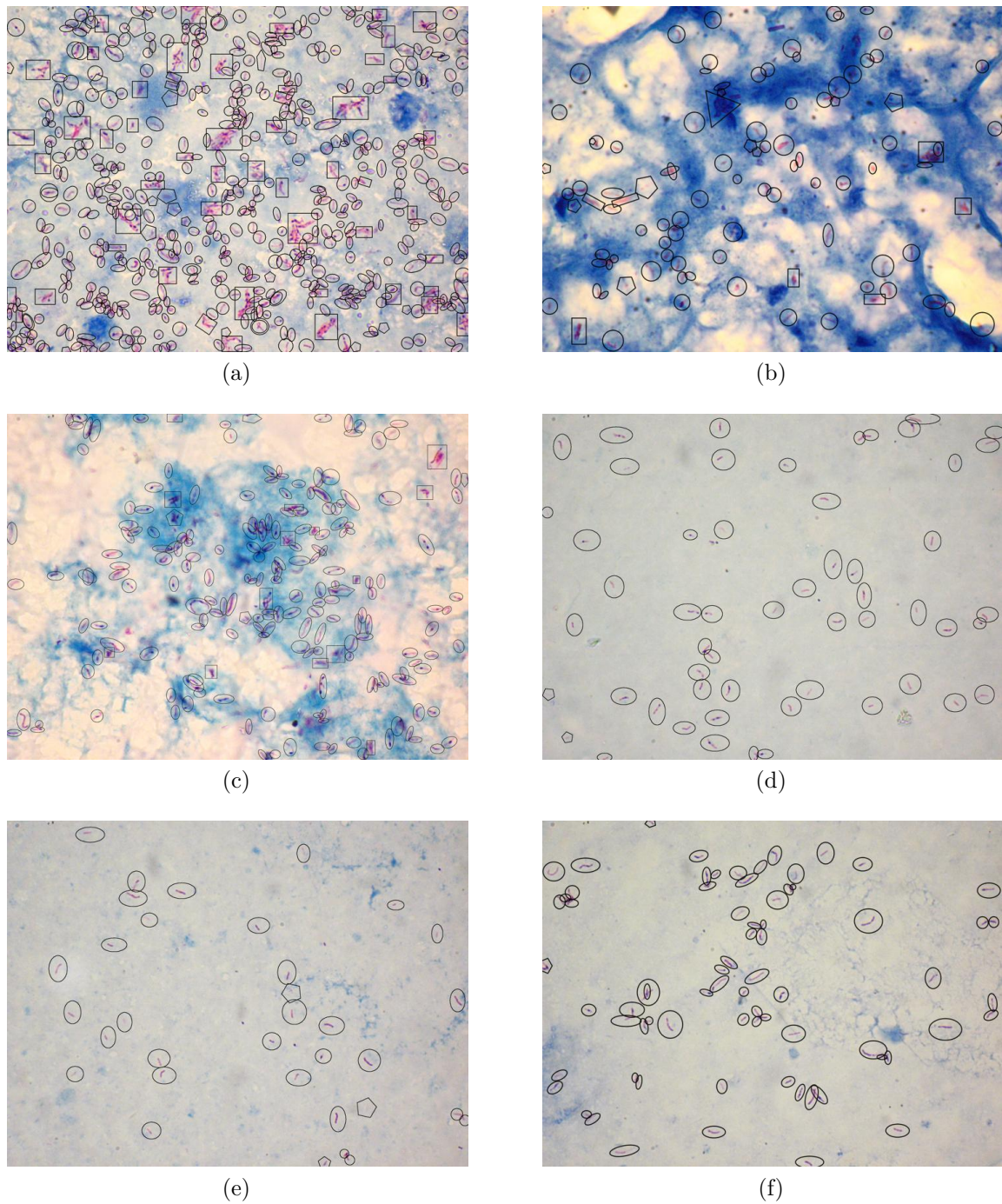


Figura 9 – Exemplos de imagens do banco de dados marcadas pelo especialista.

ser facilmente descartados. Dessa forma, é importante considerar apenas os *pixels* de fundo que possuem uma matiz similar à matiz característica dos bacilos. A Figura 10 mostra as cores distribuídas ao longo do intervalo normalizado de matiz.

Essa seleção é importante para garantir que os *pixels* de fundo possuam uma única distribuição de probabilidade, para que a análise do discriminante linear de Fisher possa ser aplicada.

Nesta análise foram utilizadas 60 imagens, sendo 5 de cada lâmina, nas quais os

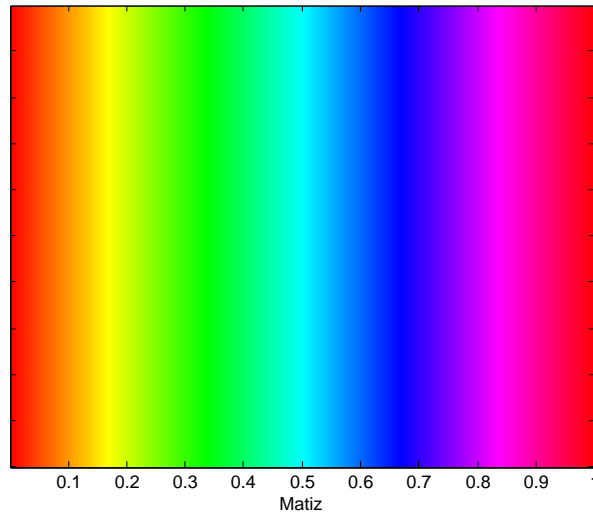


Figura 10 – Distribuição de cores conforme o matiz.

bacilos identificados pelo especialista foram marcados manualmente. Os bacilos identificados como duvidosos pelo especialista não foram incluídos em nenhuma das duas classes de *pixels* de bacilo e de *pixels* de fundo.

O vetor de projeção possui norma unitária, mas o valor do *pixel* projetado pode ser maior do que 1, já que o vetor do *pixel* da imagem colorida pode ter norma maior do que 1. Nesse caso, é preciso fazer a seguinte normalização

$$f_n(x, y) = \frac{f(x, y) - \min}{\max - \min}, \quad (3.1)$$

onde $f_n(x, y)$ é o valor do *pixel* normalizado, $f(x, y)$ é o valor do *pixel* não normalizado, \min é o menor valor dos *pixels* da imagem não normalizada e \max é o maior valor dos *pixels* da imagem não normalizada.

A Equação (3.1) permite incluir todas as imagens em um mesmo intervalo para que um único método de processamento possa ser aplicado a todas elas.

Dois métodos foram desenvolvidos para a segmentação: um baseado em limiarização global, cujos princípios foram apresentados na Seção 2.2.1; e outro baseado no método de Otsu, cuja fundamentação foi apresentada na Seção 2.2.2.

3.2.1.1 Segmentação baseada em limiarização global

Conforme discutido na Seção 2.2.1, a segmentação por limiarização global necessita da especificação de um limiar T . Neste trabalho, o limiar de segmentação é calculado a partir da intensidade média dos *pixels* da imagem projetada. A motivação para a utilização de um processo de segmentação por limiarização global é sua simplicidade e seu baixo custo computacional. Além disso, espera-se que nas imagens mais escuras, onde a intensidade média dos *pixels* é menor e, portanto, um limiar menor deve ser utilizado de modo a incluir

os bacilos; e em imagens mais claras, cuja intensidade média dos *pixels* é maior, um limiar maior deve ser usado para o mesmo fim. Espera-se, portanto, que o limiar de segmentação seja proporcional à intensidade média dos *pixels* da imagem projetada.

Dessa maneira, nesse trabalho buscou-se relacionar a intensidade média dos *pixels* da imagem projetada com o limiar de segmentação para que fosse possível ao sistema desenvolvido decidir o limiar a partir das informações contidas na imagem projetada. No conjunto de imagens utilizadas para a análise, o limiar a ser relacionado com a intensidade média é definido como o nível de cinza onde os histogramas dos *pixels* de bacilo e de fundo se interceptam, como mostrado na Figura 11.

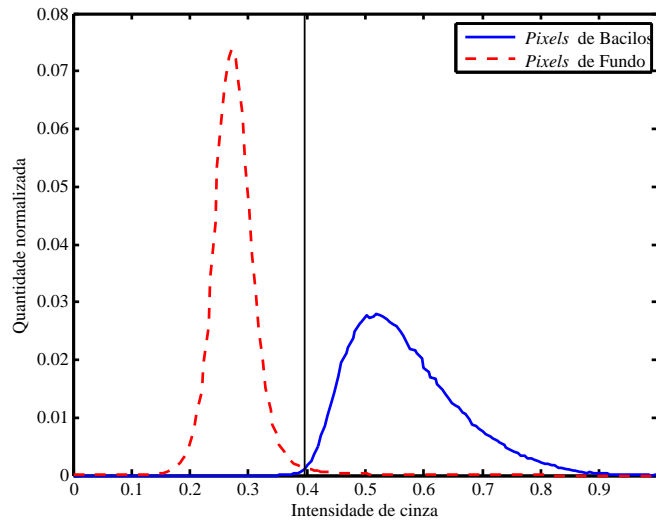


Figura 11 – Nível de cinza utilizado para a escolha do limiar.

A Figura 12 mostra a relação entre o nível de cinza no qual os histogramas se interceptam das 60 imagens utilizadas para a análise e a intensidade média da imagem projetada considerando todos os *pixels*. Observa-se a proporcionalidade direta entre uma variável e outra. Além disso, a relação entre as variáveis parece ser bem representada por uma função de primeiro grau.

O polinômio de primeiro grau obtido a partir da regressão linear é

$$T = 0,8363I_{média} + 0,1444 \quad (3.2)$$

onde T é o limiar e $I_{média}$ é a intensidade média dos *pixels*. A Figura 13 mostra o polinômio de primeira ordem obtido por regressão linear a partir dos pontos mostrados na Figura 12.

O processo de segmentação por limiarização global consiste então em, a partir da imagem projetada conforme a metodologia descrita na Seção 3.2.1, encontrar a intensidade média dos *pixels* da imagem e determinar pela Equação (3.2) o limiar T descrito na Seção 2.2.1.

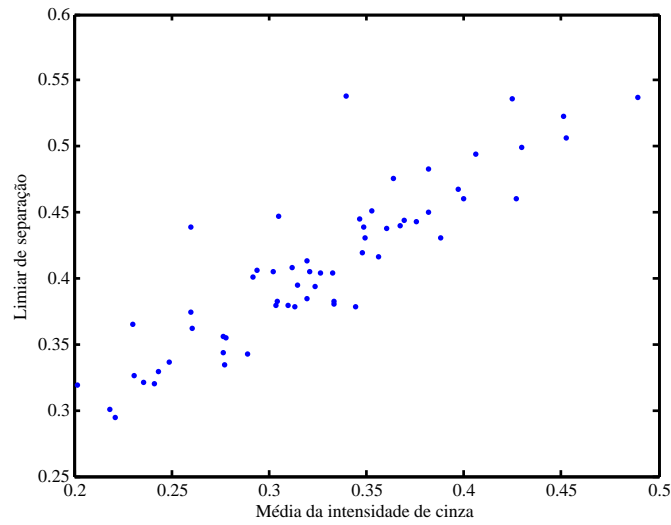


Figura 12 – Relação entre o limiar escolhido e a intensidade média das imagens projetadas.

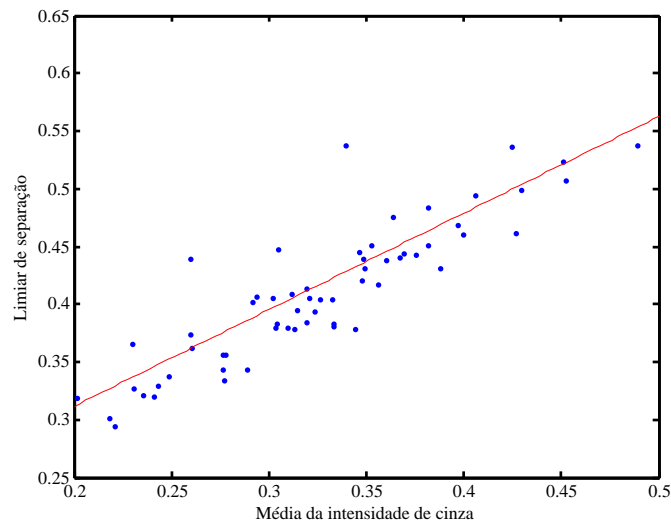


Figura 13 – Função de primeiro grau obtido por regressão linear da relação entre o limiar escolhido e a intensidade média das imagens projetadas.

3.2.1.2 Eliminação de áreas

Após a segmentação, podem restar ainda estruturas com áreas muito grandes ou muito pequenas em comparação com as estruturas que, de fato, representam bacilos de tuberculose.

Além do ruído, que contribui para o aparecimento de ambos os tipos de estruturas, estruturas com áreas muito pequenas também são geradas por bacilos que estão fora de foco. A solução deste problema representa um desafio no desenvolvimento de um sistema de diagnóstico automático de tuberculose, mas está fora do escopo deste trabalho. Estruturas com áreas muito grandes também são geradas por agrupamentos de bacilos que aparecem

em imagens que contém muitos bacilos.

O ruído neste caso é causado pelo processo de compressão com perdas das imagens em formato JPEG, por pequenas variações na iluminação das imagens e ainda pela coloração não-uniforme que as estruturas microscópicas presentes na lâmina apresentam após a aplicação da técnica de Ziehl-Neelsen. É interessante notar que variações na iluminação de imagens geralmente causam grandes problemas para o processamento digital destas. Entretanto, nas imagens deste trabalho a iluminação é muito próxima da uniforme, de forma que nenhum tratamento especial foi necessário para corrigir os problemas associados à iluminação.

As estruturas com áreas muito pequenas podem ser eliminadas, pois não representam bacilos, ou, no caso de bacilos fora de foco, apenas parte do bacilo é segmentado, de modo que a forma da estrutura não representa a forma de um bacilo corretamente segmentado. As estruturas com áreas muito grandes também podem ser eliminadas. Neste caso, como os agrupamentos muito grandes de bacilos aparecem apenas em imagens que contém muitas bactérias, mesmo que os agrupamentos sejam eliminados, ainda será possível identificar em grande quantidade bacilos isolados na imagem.

A Figura 14 mostra o histograma das áreas das estruturas que representam bacilos nas 60 imagens utilizadas nesta análise. Na Figura 14 estão incluídos bacilos isolados e agrupamentos de bacilos.

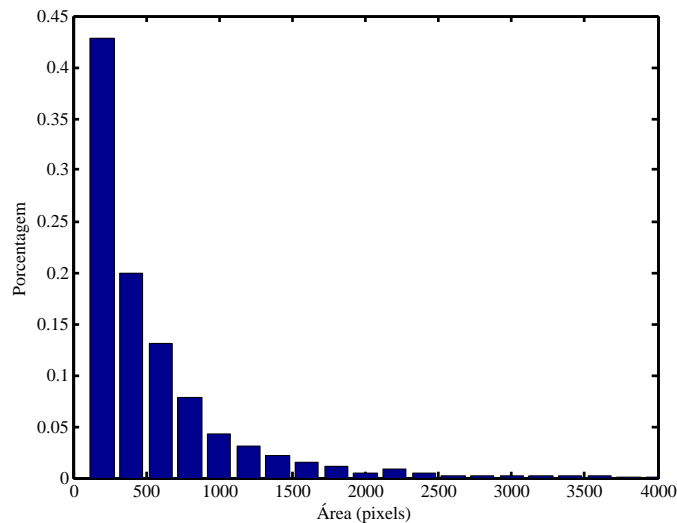


Figura 14 – Histograma das áreas de bacilos.

Com base no histograma mostrado na Figura 14, o intervalo escolhido para as áreas das estruturas que seguirão para a próxima etapa é $[200 - 2.600]$ *pixels*. Enquanto o valor para o limite inferior de 200 *pixels* foi o mesmo utilizado em Costa et al. (2008), o valor para o limite superior inclui mais de 95% dos bacilos e agrupamentos de bacilos presentes

nas 60 imagens utilizadas para análise, de modo que cinco imagens de cada paciente foram usadas para a análise.

3.2.1.3 Pós-processamento morfológico

Finalmente, a última etapa da segmentação das imagens baseada em limiarização global é fazer, sobre a imagem segmentada, uma operação de fechamento seguida de uma operação de abertura. Essa etapa busca suavizar as estruturas segmentadas aproveitando as vantagens de cada uma das operações, como detalhado na Seção 2.3.

O elemento estruturante utilizado foi o círculo com raio igual a 2 *pixels*. Essa escolha foi empírica e baseada no tamanho das estruturas segmentadas. A utilização de um raio maior resulta em uma grande deformação das estruturas, enquanto o círculo foi escolhido por sua condição de simetria.

O objetivo neste caso é de fazer apenas uma suavização superficial isométrica nos contornos das estruturas, o que justifica a escolha da forma e do raio utilizados. A Figura 15 mostra o elemento estruturante utilizado.

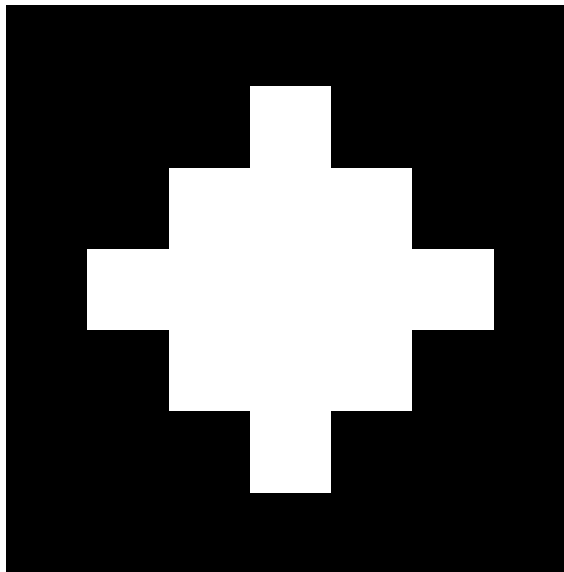


Figura 15 – Elemento estruturante na forma de um disco com raio de 2 *pixels*.

3.2.1.4 Segmentação baseada no método de Otsu

A segmentação pelo método de Otsu conforme descrito na Seção 2.2.2 sempre resulta em duas classes quando da maximização do critério apresentado na Equação (2.13). Quando da aplicação do método de Otsu em uma imagem que contém poucos bacilos, a quantidade de *pixels* de bacilos na imagem não é suficiente para maximizar o critério da Equação (2.13), de modo que o limiar de segmentação não é adequado para a segmentação dos bacilos, segmentando uma grande quantidade de *pixels* de fundo.

A solução neste caso é dividir a imagem em várias pequenas regiões, chamadas aqui de janelas devido a sua forma retangular, e implementar o método de Otsu em cada uma delas. O tamanho da janela deve ser tal que a quantidade de *pixels* de um único bacilo, na média, seja suficiente para maximizar o critério do método.

Entretanto, uma janela muito pequena pode fazer com que o inverso ocorra: a quantidade de *pixels* de fundo na janela é muito pequena de modo que o limiar de segmentação obtido pelo método de Otsu acaba excluindo alguns *pixels* de bacilo na imagem segmentada.

O tamanho da janela a ser utilizada foi determinada de forma empírica através de constatação visual dos melhores resultados a partir das imagens disponíveis. Para reduzir as possibilidades, foram avaliadas janelas cujo tamanho representavam divisões por potências de dois do tamanho da imagem original. O tamanho da janela escolhido foi então igual a $\frac{1}{16}$ do tamanho original da imagem, ou, de forma equivalente, uma janela de 132 x 176 *pixels*.

As imagens geradas são imagens binárias nas quais os bacilos isolados estão envoltos em regiões com *pixel* igual a zero e outras regiões onde não foram encontrados bacilos com muito ruído. A Figura 16 mostra o resultado da segmentação baseada no método de Otsu. Na imagem, os bacilos identificados pelo algoritmo de segmentação são mostrados como estruturas brancas isolados em uma região de fundo escuro. O ruído está presente em conjuntos de janelas onde não existem bacilos e aparecem como regiões nas quais não é possível identificar qualquer estrutura.

A remoção do ruído é feito a partir de uma operação morfológica de fechamento. O elemento estruturante utilizado, neste caso, foi o círculo com raio igual a 6 *pixels*. Esse parâmetro também foi escolhido de forma empírica através de constatação visual.

Procede-se então com a mesma eliminação de áreas e o pós-processamento morfológico usado após a segmentação baseada em limiarização global.

3.2.2 Classificação das estruturas

A segunda parte do sistema proposto consiste em classificar as estruturas que resultam das imagens segmentadas. Como estruturas a serem classificadas foram considerados os objetos 8-conectados.

Para esse fim, foram considerados dois tipos de classificadores: as redes neurais, descritas na Seção 2.5.1 e as máquinas de vetores de suporte, descritas na Seção 2.5.2. As características consideradas nos classificadores foram a excentricidade, a circularidade, os sete momentos invariantes, os cinco primeiros descritores de Fourier e a média e o desvio padrão de cada uma das componentes de cor do espaço de cores RGB. Essas características foram descritas na Seção 2.4.

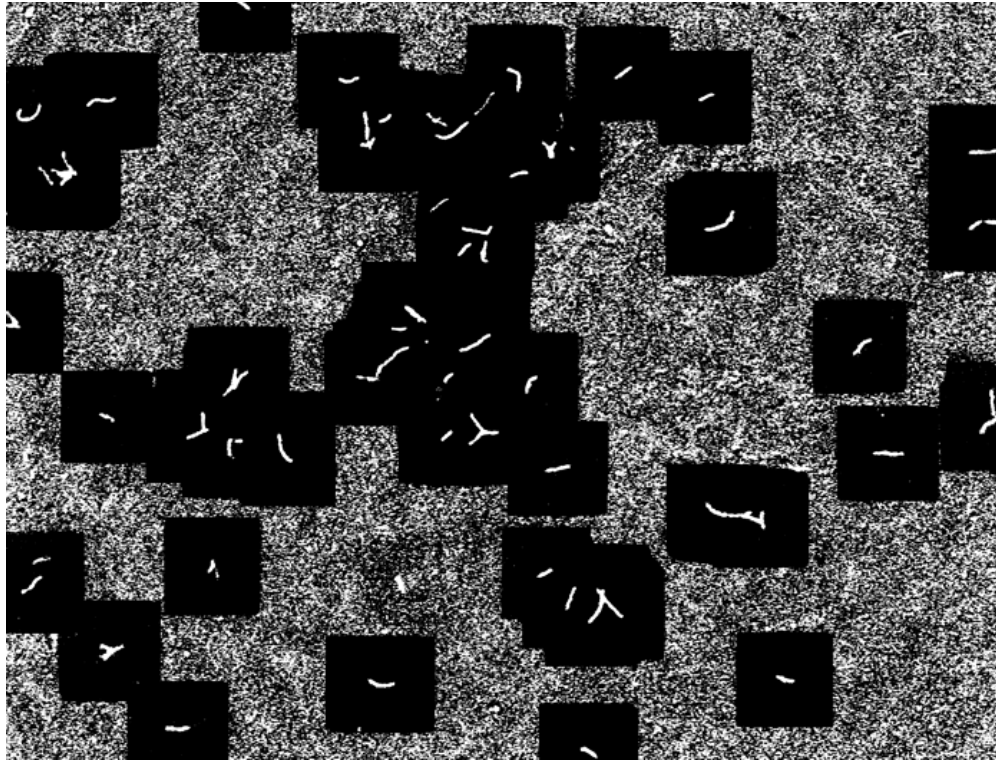


Figura 16 – Imagem gerada pela segmentação baseada no método de Otsu.

Para a avaliação dos resultados, o método da validação cruzada foi utilizada considerando que, em cada *fold*, 90% das imagens eram utilizadas para o treinamento dos classificadores e 10% das imagens eram utilizadas para o teste dos classificadores. O treinamento dos classificadores foi feito usando 10 conjuntos de treinamento e teste diferentes. Esse procedimento é conhecido na literatura como *10-fold*. A validação cruzada evita a escolha de conjuntos de treinamento e teste que não sejam representativos. O procedimento de *10-fold* é a técnica de avaliação padrão em situações onde os dados são limitados (WITTEN; FRANK, 2005).

Nesse caso, como o banco de imagens possui 120 imagens, sendo 10 de cada um dos 12 pacientes, os conjuntos de teste eram formados por 12 imagens, sendo uma imagem de cada paciente, enquanto o conjunto de treinamento continha as 108 imagens restantes, sendo 9 de cada paciente.

As redes neurais foram treinadas pelo algoritmo de retropropagação de Levenberg-Marquardt (HAGAN; MENHAJ, 1994). A taxa de aprendizado e o fator de incremento/decremento da taxa de aprendizado utilizados foram, respectivamente, 0,001 e 10, como recomendado pelo *software* MATLAB por representar um bom compromisso entre estabilidade do algoritmo e velocidade de convergência. A arquitetura adotada foi a rede neural alimentada diretamente com uma única camada oculta. Diversas topologias, no que diz respeito à quantidade de neurônios na camada oculta, foram avaliadas visando determinar o número de neurônios na camada escondida que apresentasse uma representação

adequada das características das estruturas para sua classificação. O número de neurônios escolhido para a camada oculta foi igual a 20.

Com relação às máquinas de vetores suporte, foram avaliadas as funções de *kernel* de base polinomial e de base radial, sendo a escolhida para a classificação a função polinomial de grau 5 por apresentar melhores resultados.

A avaliação da classificação é feita comparando os parâmetros de sensibilidade, especificidade, acurácia e precisão, como definido em (FAWCETT, 2006) Esses parâmetros são calculados por imagem.

No contexto de duas classes, a sensibilidade é a taxa de verdadeiros positivos em relação ao número total de positivos. No presente objetivo de classificação de estruturas segmentadas, os positivos representam todas as estruturas que pertencem a bacilos e os verdadeiros positivos são as estruturas corretamente classificadas como bacilos. A sensibilidade é matematicamente descrita como:

$$sensibilidade = \frac{TP}{TP + FN}. \quad (3.3)$$

A especificidade é a taxa de verdadeiros negativos em relação ao número total de negativos. No contexto presente, os negativos representam as estruturas segmentadas que não são bacilos e os verdadeiros negativos são as estruturas corretamente classificadas como não-bacilos. A especificidade é matematicamente descrita como:

$$especificidade = \frac{TN}{FP + TN}. \quad (3.4)$$

A acurácia é a relação entre as estruturas corretamente classificadas e o número total de estruturas:

$$acurácia = \frac{TP + TN}{P + N}. \quad (3.5)$$

Finalmente, a precisão é a relação entre os verdadeiros positivos e o número de positivos obtidos pelo classificador. Matematicamente:

$$precisão = \frac{TP}{FP + TP}. \quad (3.6)$$

Onde P é a quantidade total de positivos, N é a quantidade total de negativos, TP é o número de verdadeiros positivos, FP é o número de falsos positivos, TN é o número de verdadeiros negativos e FN é o número de falsos negativos. É interessante notar que, na prática, quanto maior o valor dessas métricas, melhor é o desempenho do classificador. Idealmente, todas elas deveriam ser 100%.

4 Resultados

4.1 Projeção das imagens

O vetor de projeção normalizado obtido pela análise do discriminante linear é mostrado na Equação 4.1.

$$w = [0,5450 \quad -0,8327 \quad 0,0984]^t. \quad (4.1)$$

A Figura 17 mostra a projeção das imagens apresentadas na Figura 8 segundo o método da análise de discriminante linear.

Observa-se que os *pixels* de bacilos apresentam uma intensidade de cinza maior (*pixels* mais claros) em comparação aos *pixels* do fundo. Essa é uma escolha feita *a priori* e o inverso (*pixels* de bacilos com uma menor intensidade de cinza em comparação aos *pixels* de fundo) também poderia ser obtido. Isso se deve ao fato de que é possível obter um vetor de projeção capaz de colocar os *pixels* de bacilos mais próximos de 0 e os *pixels* de fundo mais próximos de 1 no espaço projetado e que resulte no mesmo valor para o critério de Fisher apresentado na Equação 2.1.

A escolha do vetor de projeção que resulta nas imagens mostradas na Figura 17 visa a colaborar com o processo de segmentação subsequente ao colocar os *pixels* de bacilos mais claros do que os *pixels* de fundo.

4.2 Segmentação das imagens

4.2.1 Segmentação baseada em limiarização global

A Figura 18 mostra o resultado da segmentação usando o método detalhado na Seção 2.2.1 sobre as imagens apresentadas na Figura 17, sem a eliminação de áreas e o pós-processamento morfológico. É possível observar a existência de ruído na forma de pequenas estruturas, principalmente no canto direito das imagens, e grandes estruturas que podem ser formados por agrupamentos de bacilos, mas também por aglomerado de outras estruturas presentes nas imagens e ruído.

O resultado da segmentação foi avaliado nas 60 imagens que não foram utilizadas na determinação da função de primeiro grau apresentada na Equação 3.2 resultante da regressão linear dos dados que relacionam a intensidade média da imagem projetada com o limiar utilizado para a segmentação. 98,69% dos bacilos presentes nas 60 imagens de teste foram segmentados, de modo que apenas 1,31% dos bacilos foram perdidos.

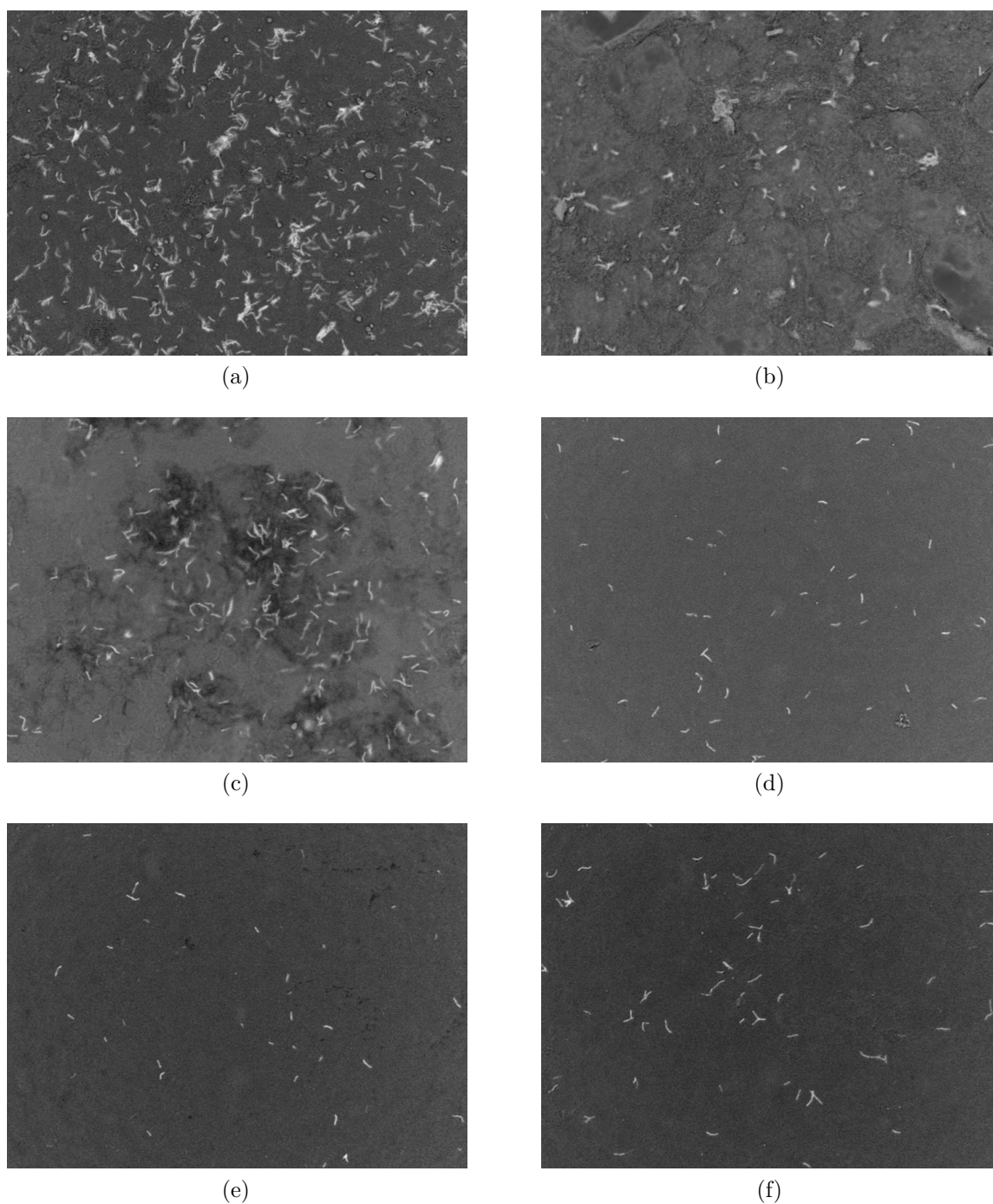


Figura 17 – Projeção das imagens mostradas na Figura 8.

A Figura 19 mostra as imagens da Figura 18 após a eliminação de áreas e o pós-processamento morfológico. Nota-se que as pequenas estruturas causados pelo ruído bem como as grandes estruturas não estão mais presentes.

Nesta etapa, 85,61% dos bacilos presentes no conjunto de 60 imagens utilizadas para avaliação do resultado da segmentação permanecem na imagem segmentada, ou seja, 14,39% dos bacilos foram perdidos nas etapas de segmentação por limiarização global, eliminação de áreas e pós-processamento morfológico.

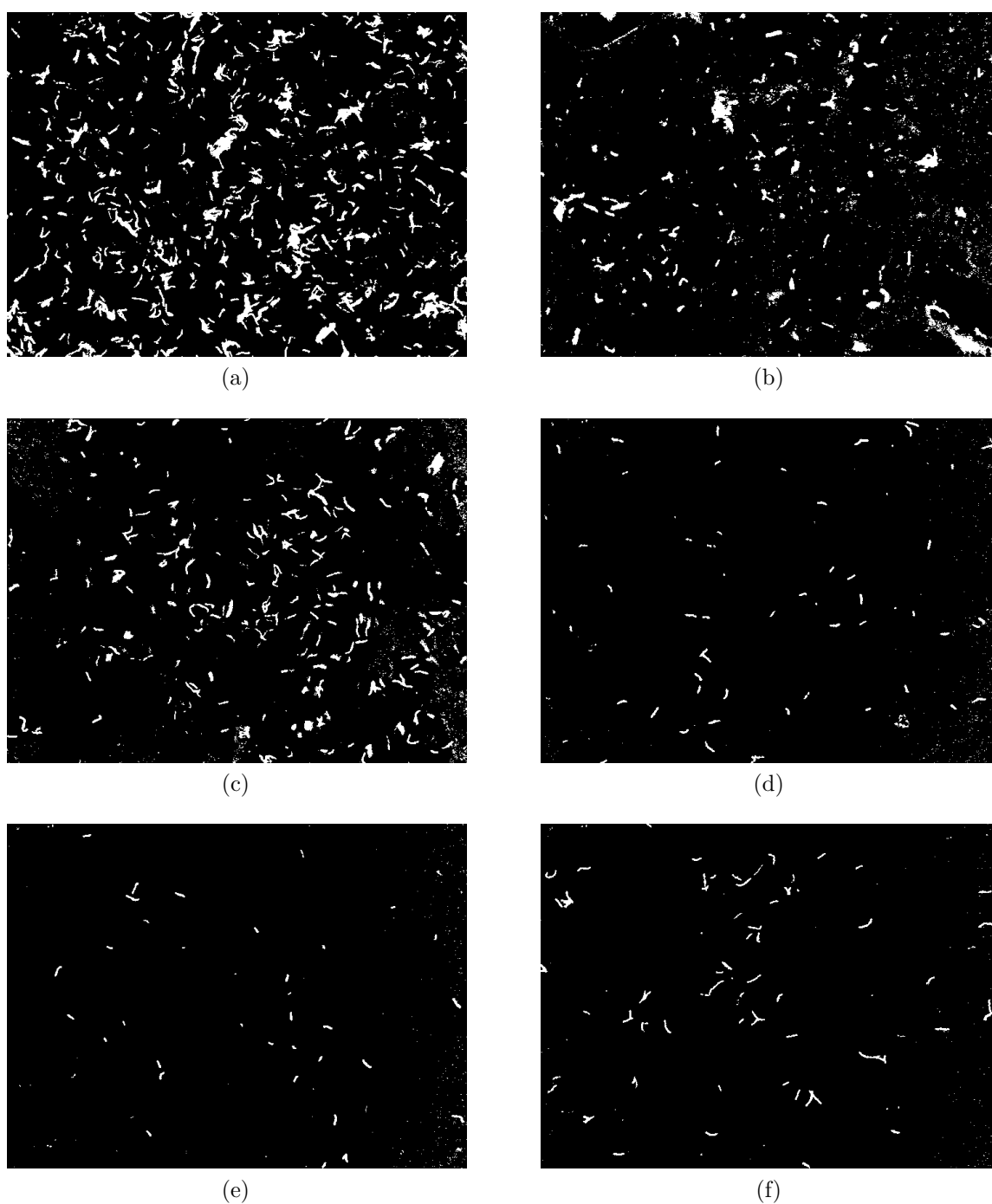


Figura 18 – Segmentação das imagens mostradas na Figura 17 pelo método de limiarização global.

4.2.2 Segmentação baseada no método de Otsu

O resultado da segmentação das imagens mostradas na Figura 17 pela técnica baseada no método de Otsu apresentada na Seção 2.2.2 é apresentada na Figura 20. É possível observar bacilos isolados e regiões com ruído intenso. Isso se deve à própria característica do método de Otsu que sempre encontra algum limiar de segmentação, de modo que em todas as janelas existirão sempre duas classes. Nas janelas em que existem

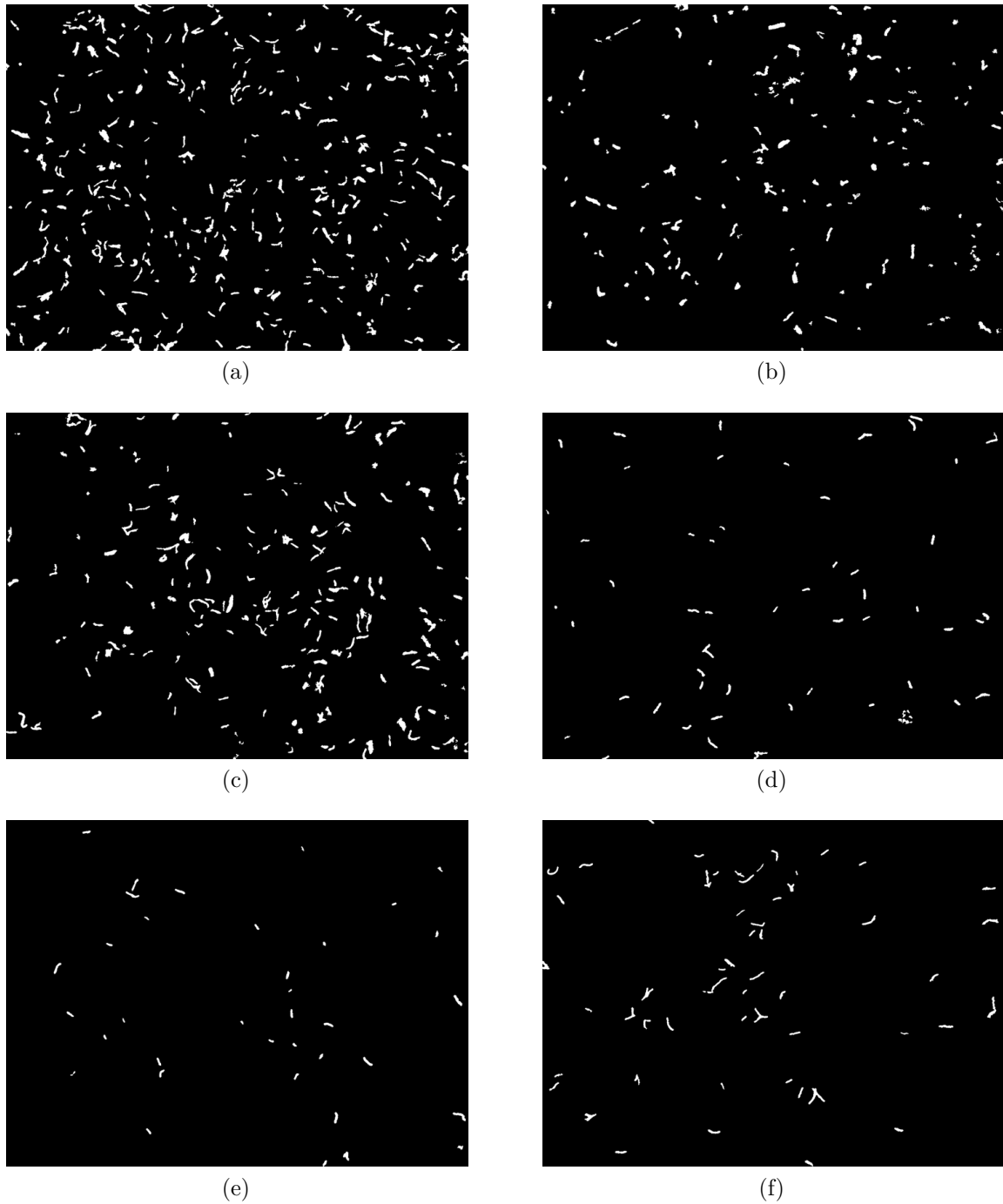


Figura 19 – Resultado da eliminação de áreas muito grandes e muito pequenas e pós-processamento morfológico das imagens mostradas na Figura 18.

bacilos, uma das classes é o próprio bacilo na maioria das vezes. Mas em janelas onde não há bacilos, a segmentação separa *pixels* pertencentes à mesma classe (*pixels* de fundo) gerando, de forma errônea, duas classes, quando, na verdade, nenhum *pixel* deveria ter sido segmentado. Esse processo aumenta o nível de ruído, levando aos resultados mostrados na Figura 20, onde é possível notar que as imagens com maior quantidade de bacilos, (a), (b) e (c), apresentam uma menor quantidade de ruído em comparação com as imagens com menor quantidade de bacilos, (d), (e) e (f).

Mesmo nas áreas com ruído intenso podem existir bacilos cuja área não foi grande o suficiente para que fossem corretamente isolados dos *pixels* de fundo pelo método de Otsu e, portanto, não aparecem como bacilos isolados.

Nesta etapa, todos os bacilos existentes nas 60 imagens usadas para avaliação foram segmentados.

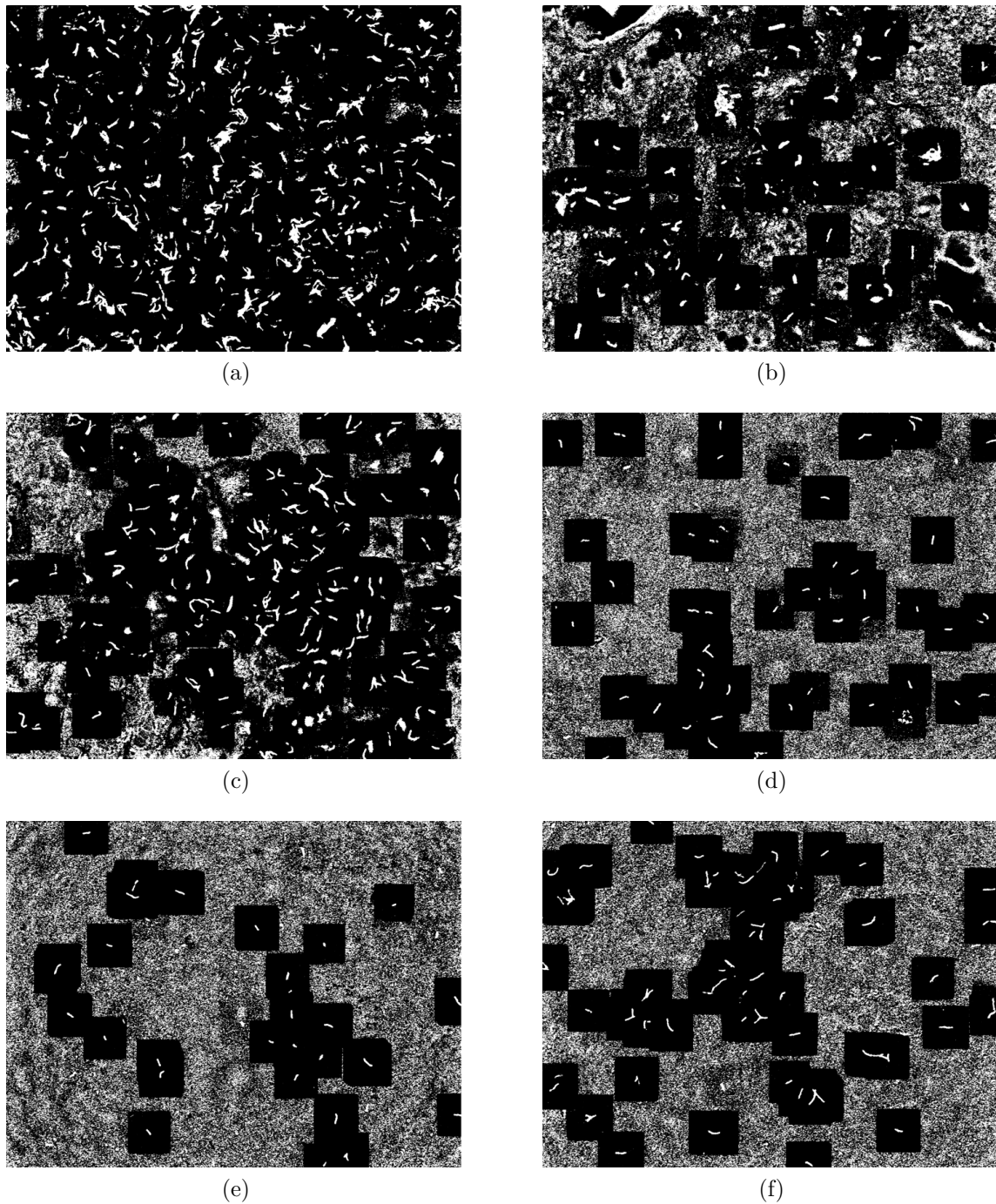


Figura 20 – Segmentação das imagens mostradas na Figura 17 pelo Método de Otsu modificado.

A Figura 21 apresenta as imagens mostradas na Figura 20 após a operação de

fechamento, a eliminação de áreas e o pós-processamento morfológico. Nota-se que o ruído é eliminado assim como as grandes estruturas.

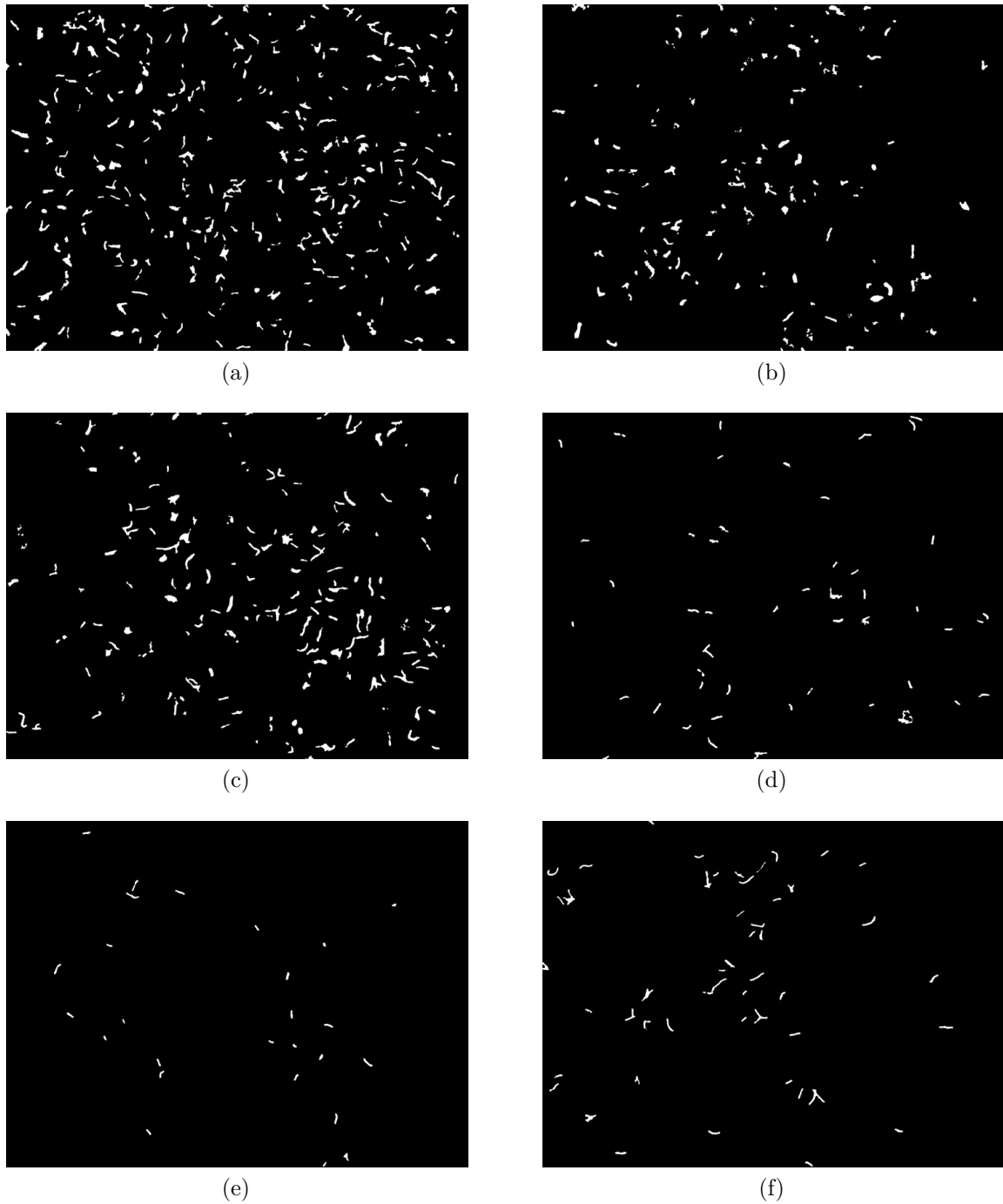


Figura 21 – Resultado da eliminação de áreas muito grandes e muito pequenas e pós-processamento morfológico das imagens mostradas na Figura 20.

Nesta etapa, 76,34% dos bacilos presentes no conjunto de 60 imagens utilizadas para avaliação do resultado da segmentação permanecem na imagem segmentada, ou seja, 23,66% dos bacilos foram perdidos nas etapas de segmentação baseada no método de Otsu mostrado na Seção 2.2.2, eliminação de áreas e pós-processamento morfológico. É

interessante notar que a maior perda ocorre nas imagens com muitos bacilos, e a maior eliminação é dos grandes agrupamentos de bacilos.

4.3 Classificação das estruturas

4.3.1 Classificação usando redes neurais

As tabelas contendo os resultados da classificação das estruturas segmentadas usando redes neurais para cada um dos *folds* é apresentada na Seção B.1 do Apêndice B.

A Tabela 1 resume os valores de sensibilidade, especificidade, precisão e acurácia para cada um dos *folds* da classificação das estruturas segmentadas.

Tabela 1 – Resumo dos valores de média e desvio padrão para cada *fold* da classificação das estruturas segmentadas usando redes neurais.

<i>FOLD</i>	SENSIBILIDADE		ESPECIFICIDADE		PRECISÃO		ACURÁCIA	
	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO
1	91,12%	14,68%	82,51%	29,50%	92,99%	6,94%	89,43%	15,38%
2	79,94%	19,09%	92,64%	13,95%	92,31%	8,95%	89,14%	15,29%
3	94,25%	6,44%	91,43%	12,99%	94,29%	5,96%	89,34%	12,56%
4	88,15%	14,60%	82,37%	29,35%	93,99%	5,70%	88,30%	18,93%
5	86,79%	21,81%	84,93%	29,31%	95,73%	4,38%	85,97%	17,91%
6	83,15%	27,66%	83,21%	28,67%	93,31%	4,55%	85,76%	24,41%
7	83,64%	20,12%	83,40%	29,13%	93,36%	6,88%	91,41%	12,53%
8	85,31%	19,70%	83,94%	9,35%	94,90%	6,26%	92,50%	14,15%
9	91,60%	10,42%	90,16%	15,51%	95,36%	3,92%	89,05%	15,90%
10	84,50%	19,64%	94,71%	9,17%	93,93%	7,11%	81,25%	21,62%

4.3.2 Classificação usando máquinas de vetores suporte

As tabelas contendo os resultados da classificação das estruturas segmentadas usando redes neurais para cada um dos *folds* é apresentada na Seção B.2 do Apêndice B.

A Tabela 2 resume os valores de sensibilidade, especificidade, precisão e acurácia para cada um dos *folds* da classificação das estruturas segmentadas.

Tabela 2 – Resumo dos valores de média e desvio padrão para cada *fold* da classificação das estruturas segmentadas usando máquinas de vetores suporte.

<i>FOLD</i>	SENSIBILIDADE		ESPECIFICIDADE		PRECISÃO		ACURÁCIA	
	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO
1	89,57%	15,12%	85,31%	29,28%	93,38%	7,49%	89,78%	16,64%
2	79,23%	19,77%	93,13%	12,91%	92,42%	8,25%	85,68%	22,13%
3	93,35%	8,44%	91,33%	12,34%	94,03%	5,74%	86,28%	18,48%
4	87,31%	14,09%	82,36%	29,28%	93,45%	5,72%	89,34%	17,30%
5	87,60%	18,63%	84,71%	29,48%	95,70%	4,32%	86,12%	18,09%
6	82,75%	28,68%	83,39%	29,00%	94,30%	4,70%	86,89%	23,50%
7	83,83%	20,57%	93,32%	29,12%	93,60%	6,48%	91,84%	11,66%
8	82,38%	20,90%	93,34%	9,76%	94,26%	8,43%	89,20%	14,96%
9	90,21%	13,55%	90,20%	15,72%	95,45%	4,19%	88,94%	15,21%
10	86,18%	17,94%	95,33%	8,11%	94,96%	6,48%	81,11%	22,91%

5 Análise dos resultados

5.1 Projeção das imagens

A projeção das imagens é computacionalmente eficiente, uma vez que para um único *pixel* são necessárias apenas três operações de multiplicação e duas operações de soma. Dessa forma, para uma imagem com n *pixels*, o esforço computacional da projeção é da ordem de $O(5n)$.

A Figura 22 apresenta os histogramas das imagens mostradas na Figura 17 e mostra a separação entre as duas classes, maximizada no sentido do critério de Fisher como apresentado na Equação (2.1).

5.2 Segmentação das imagens

A perda de bacilos no processo de segmentação baseada em limiarização global se deve aos bacilos que ficaram abaixo do limiar estabelecido pela função de primeiro grau apresentada na Equação (3.2). Como mostrado na Seção 4.2.1, apenas 1,31% dos bacilos ficaram abaixo do limiar mencionado o que sugere que o método é adequado e eficiente na segmentação dos bacilos das imagens projetadas conforme o método da análise do discriminante linear.

Embora a segmentação baseada no método de Otsu tenha segmentado 100% dos bacilos, essa é uma característica inerente à técnica utilizada, uma vez que os bacilos são sempre mais claros do que o fundo, os bacilos são sempre segmentados, ainda que vários *pixels* de fundo sejam segmentados junto e impeçam a distinção entre os *pixels* de bacilo e os *pixels* de fundo.

Com relação ao esforço computacional, para uma imagem de n *pixels*, o custo computacional do método de segmentação baseada em limiarização global é da ordem de $O(2n)$ devido ao cálculo da média dos *pixels* da imagem, o cálculo do limiar de segmentação e a segmentação por limiarização global em si.

O método de Otsu possui um custo computacional da ordem de $O(256n)$, para uma imagem de 8 bits. Como o método de segmentação baseado no método de Otsu executa a técnica de Otsu para cada janela, o custo computacional é proporcional ao número de janelas analisadas por imagem. Neste caso, o esforço computacional é da ordem de $O(10^9n)$, muito superior à técnica de segmentação baseada em limiarização global.

Dessa forma, o método baseado em limiarização global é preferível se comparado à

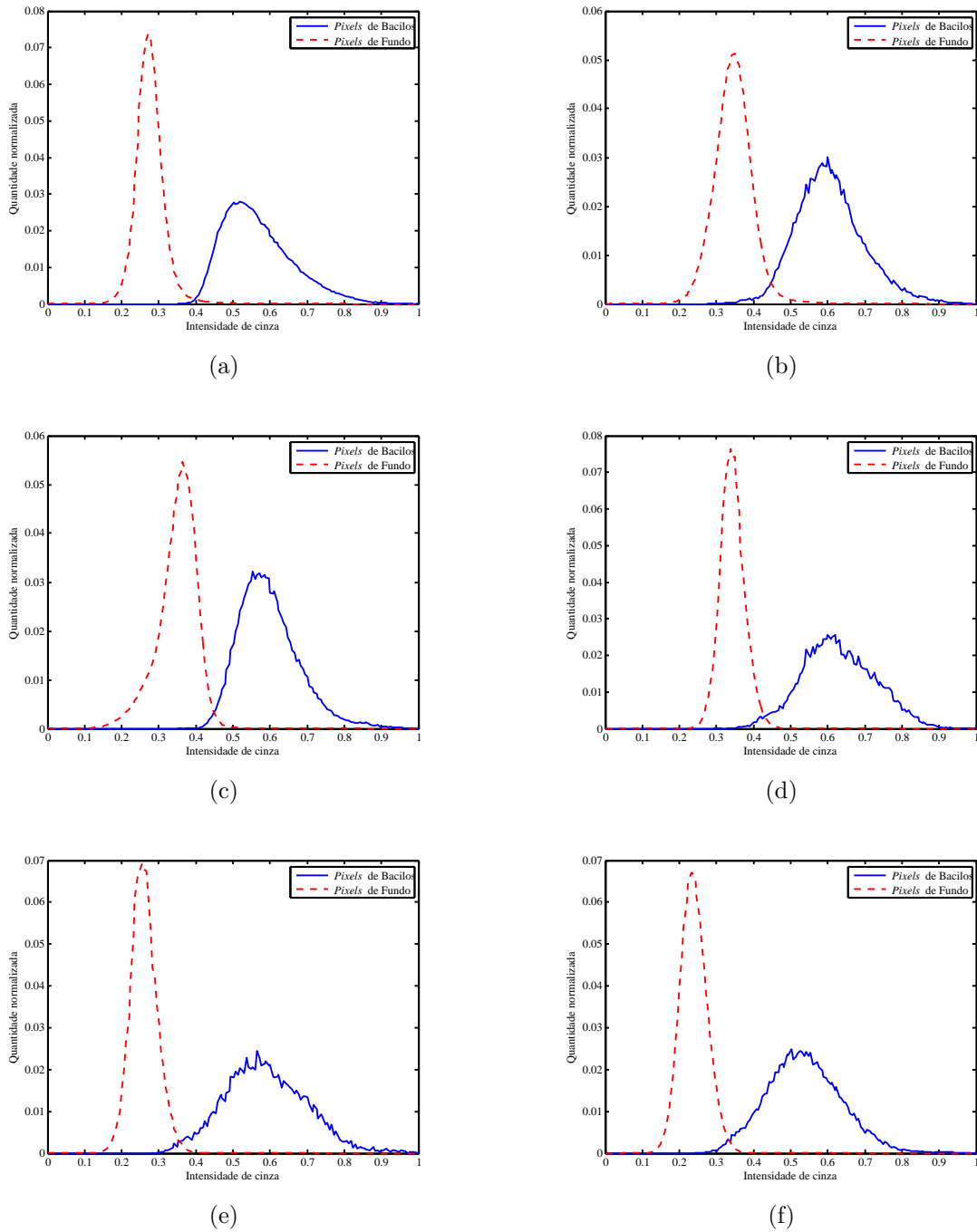


Figura 22 – Histograma da projeção das imagens mostradas na Figura 17.

técnica baseada no método de Otsu. A única vantagem do segundo método sobre o primeiro é o fato de não ser baseado nos dados do conjunto de treinamento. A única informação obtida a partir do conjunto de dados para o segundo método é o tamanho adequado da janela que vai percorrer a imagem aplicando o método de Otsu para segmentar os bacilos, conforme explicado na Seção 2.2.2.

A quantidade de bacilos perdidos após as etapas de eliminação de áreas e pós-processamento morfológico é maior devido principalmente à eliminação de estruturas com

grandes áreas, muitas delas formadas por grandes aglomerações de bacilos que formam uma única estrutura após a etapa de segmentação. Entretanto, as imagens que possuem grandes aglomerações de bacilos também possuem uma grande quantidade de bacilos isolados, de modo que a identificação de bacilos não é comprometida.

Os bacilos que, após a etapa de segmentação, formam estruturas com área muito pequena e são eliminadas se devem principalmente a bacilos que estão fora de foco nas imagens originais. Isso faz com que sua cor seja diferente da coloração rosa-avermelhada característica da técnica de coloração de Ziehl-Neelsen e, conseqüentemente, a projeção utilizada seja ineficiente sobre esses *pixels*. Nesse caso, a varredura da lâmina com o esfregaço do escarro do paciente e o ajuste do foco no processo de aquisição de imagens irá propiciar a detecção desses bacilos usando o método apresentado neste trabalho.

Embora resulte em uma perda elevada de bacilos, se comparada à segmentação, a eliminação de grandes e pequenas estruturas é importante para reduzir a variabilidade das estruturas na etapa de classificação.

5.3 Classificação das estruturas

As Tabelas 3 e 4 apresentam o resumo das médias de cada *fold*. Esses valores foram calculados a partir dos valores de média das tabelas que apresentam o resultado dos classificadores (mostradas no apêndice). As médias representam um bom indicador por levar em conta os resultados obtidos em todas as imagens, em oposição ao máximo e mínimo que representam o melhor e o pior caso, respectivamente.

Tabela 3 – Resumo das médias de cada *fold* da classificação usando redes neurais.

	MÉDIA	MÁXIMO	MÍNIMO	DESVIO PADRÃO
SENSIBILIDADE	86,84%	94,25%	79,94%	4,43%
ESPECIFICIDADE	87,93%	94,71%	82,37%	5,10%
ACURÁCIA	94,02%	95,73%	92,31%	1,08%
PRECISÃO	88,22%	92,50%	81,25%	3,20%

Tabela 4 – Resumo das médias de cada *fold* da classificação usando máquinas de vetores suporte.

	MÉDIA	MÁXIMO	MÍNIMO	DESVIO PADRÃO
SENSIBILIDADE	86,24%	93,35%	79,23%	4,25%
ESPECIFICIDADE	88,24%	95,33%	82,36%	4,91%
ACURÁCIA	94,16%	95,70%	92,42%	1,01%
PRECISÃO	87,52%	91,84%	81,11%	3,20%

Os resultados de sensibilidade usando redes neurais foram maiores do que o resultado obtido utilizando máquinas de vetores suporte, embora estejam muito próximos. De forma

geral, os resultados obtidos pelos dois classificadores foram muito próximos, o que indica que as características utilizadas representam satisfatoriamente as estruturas segmentadas, permitindo uma distinção eficiente entre as duas classes.

Analizando os resultados obtidos em cada imagem, de forma geral, observa-se que ambos os classificadores tiveram desempenho muito similar para cada uma delas, o que indica que o resultado da classificação das estruturas, neste caso, depende mais das características das estruturas segmentadas do que o tipo de classificador utilizado.

Os menores valores de sensibilidade foram, na maior parte dos casos, registrados nas imagens referentes ao sexto paciente. A Figura 23 mostra algumas dessas imagens. Pode-se observar que muitos bacilos estão fora de foco e a coloração azul é diferente das demais. Isso mostra como a correta execução da técnica de Ziehl-Neelsen é importante para o desempenho do sistema automático de reconhecimento de bacilos. Na maior parte dos casos, essas imagens são as responsáveis pelos resultados iguais a 0% nas Tabelas 7, 8, 9, 10, 13, 14, 15 e 16 do Apêndice B.

Com relação aos bacilos duvidosos, a Tabela 5 mostra a média e o desvio padrão da taxa de classificação das estruturas duvidosas como bacilos.

Tabela 5 – Média e desvio padrão da taxa de classificação de estruturas duvidosas como bacilos usando redes neurais artificiais e máquinas de vetores suporte como classificadores.

	MÉDIA	DESVIO PADRÃO
Redes Neurais Artificiais	49,00%	14,55%
Máquinas de Vetores Suporte	48,56%	15,54%

A baixa taxa de classificação das estruturas duvidosas como bacilos em ambos os classificadores é um indicador de que as características utilizadas para descrever as estruturas segmentadas são de fato representativas e se assemelham à forma como o especialista humano classifica visualmente as estruturas que encontra na lâmina.

Essa conclusão vem do fato de que, da mesma forma que o especialista humano não soube afirmar com certeza se a estrutura duvidosa era um bacilo, a taxa de classificação das estruturas segmentadas duvidosas como bacilos é próxima de 50%, o que equivale a uma escolha aleatória com probabilidade uniforme.

5.4 Comparação com outros trabalhos

Como as imagens utilizadas no presente trabalho são as mesmas utilizadas em Costa et al. (2008) e Costa Filho et al. (2012), é interessante comparar os resultados obtidos por este trabalho com os resultados obtidos por aqueles.

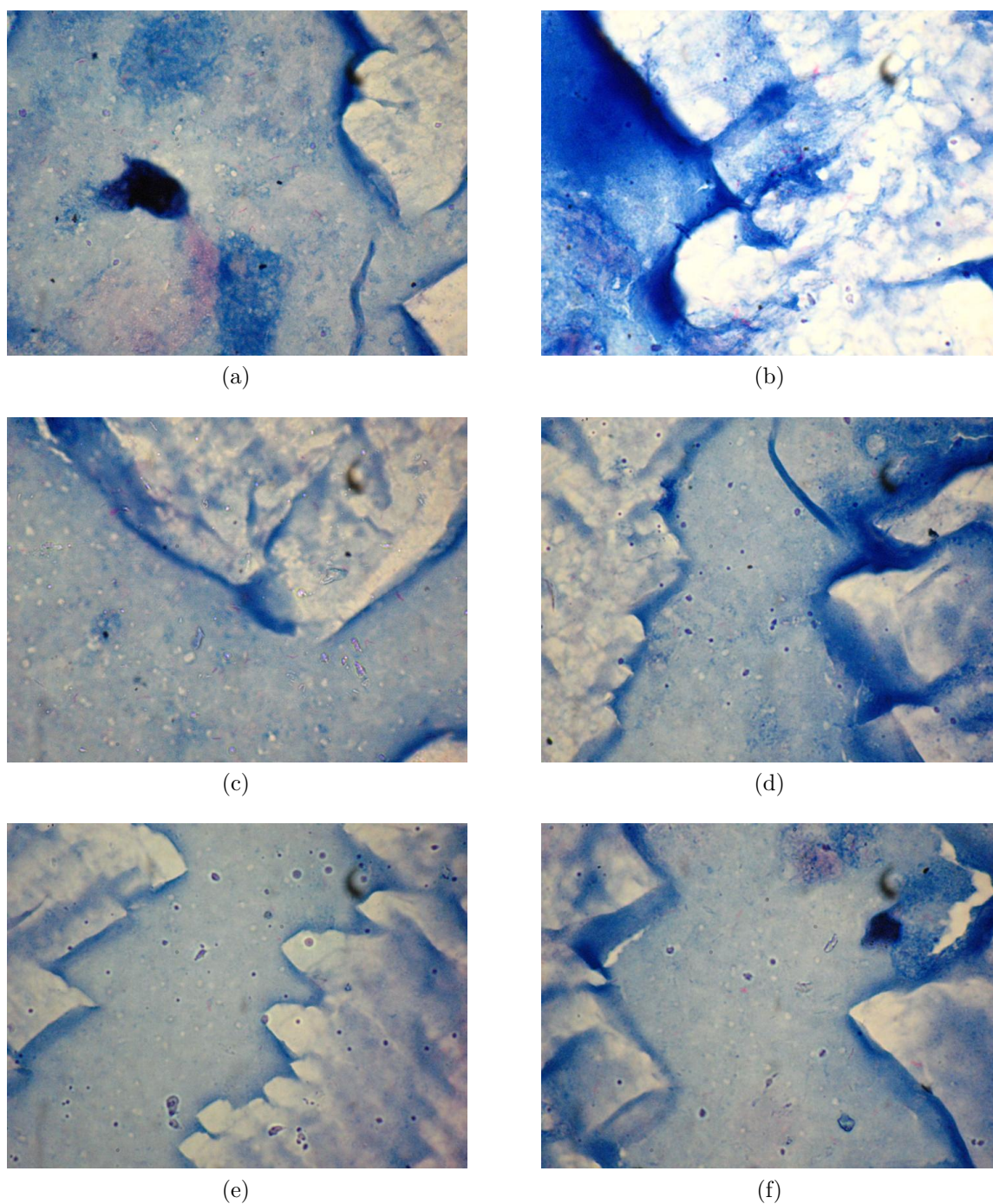


Figura 23 – Algumas das imagens relativas ao sexto paciente.

Em Costa et al. (2008), apenas a etapa de segmentação foi implementada e a melhor taxa de segmentação de bacilos foi igual a 76,65 %, enquanto que no presente trabalho, 85,61 % dos bacilos foram segmentados usando a técnica de segmentação baseada em limiarização global. Isso representa um ganho de 11,69 % do método apresentado sobre o primeiro. O esforço computacional entre os dois métodos é muito similar, de modo que nenhum ganho representativo foi obtido neste quesito.

Com relação a Costa Filho et al. (2012), a etapa de segmentação é avaliada com

relação à correta segmentação de *pixels*, enquanto no presente trabalho a avaliação da segmentação é feita em relação ao número de bacilos corretamente segmentados. Dessa forma, não é possível comparar os dois resultados. Com relação à classificação, a Tabela 6 compara os resultados obtidos neste trabalho com os resultados obtidos em Costa Filho et al. (2012).

Tabela 6 – Comparação entre a classificação de estruturas com o trabalho realizado por Costa Filho et al. (2012).

	SENSIBILIDADE	PRECISÃO	FALSA DETECÇÃO
Presente trabalho	94,25%	92,50%	7,50%
Costa Filho et al. (2012)	91,53%	91,49%	8,51%

Os critérios utilizados para a avaliação dos resultados foram a sensibilidade, a precisão e a falsa detecção, definida como $1 - \text{precisão}$. É preciso salientar que a avaliação por meio de validação cruzada não foi implementada em Costa Filho et al. (2012). Além disso, não são fornecidos detalhes sobre a avaliação das estruturas duvidosas.

Com relação à eficiência computacional, em Costa Filho et al. (2012) a segmentação é feita utilizando uma rede neural com uma camada escondida com 3 neurônios e que considera 5 características ($G - I$, $L - a$, $Y - Cr$, a , $R - G$, dos espaços de cores RGB, HSI, Lab e YCbCr). Neste caso, além do esforço computacional para representar cada um dos espaços de cores, ainda existe o custo computacional para calcular a classe de cada *pixel* pela rede neural. No presente trabalho, a projeção é feita sobre o espaço de cores RGB seguida apenas pelo cálculo da média e do limiar de segmentação, de modo que o custo computacional na etapa de segmentação é inferior àquele apresentado em Costa Filho et al. (2012).

É interessante notar que os resultados obtidos neste trabalho para a classificação das estruturas é similar àqueles obtidos em trabalhos que fizeram o reconhecimento automático de bacilos de tuberculose em imagens de microscopia por fluorescência como Veropoulos et al. (1998), Forero-Vargas et al. (2002) que, tradicionalmente, apresentam resultados superiores àqueles que utilizam microscopia convencional.

6 Conclusão

Os resultados apresentados mostram que é possível obter um bom reconhecimento automático de bacilos de tuberculose em imagens coloridas pela técnica de Ziehl-Neelsen de baixo custo computacional através da projeção das imagens baseada em análise de discriminante linear e a segmentação por limiarização global cujo limiar é definido baseado na intensidade média da imagem e a posterior classificação das estruturas segmentadas usando redes neurais artificiais ou máquinas de vetores suporte.

A técnica de segmentação baseada no método de Otsu desenvolvida mostrou-se de baixa capacidade de segmentação e alto custo computacional, tornando a sua utilização inviável.

Não existe grande diferença entre a aplicação de redes neurais artificiais ou máquinas de vetores suporte no contexto deste trabalho, o que indica que as características utilizadas para descrever as estruturas (momentos invariantes, descritores de Fourier, excentricidade, circularidade e a média e o desvio padrão de cada um dos canais de cor do espaço RGB) são adequadas para a representação dos bacilos.

O método desenvolvido, além de possuir boa capacidade de identificação de bacilos, é de baixo custo computacional, o que permite futuramente avaliar sua aplicação em sistemas especialistas de auxílio ao diagnóstico de tuberculose.

Outro fato importante sobre o sistema desenvolvido, é a capacidade de encontrar bacilos de tuberculose em imagens com poucos bacilos. Essa capacidade é essencial para um diagnóstico preciso e correto, uma vez que as imagens com poucos bacilos são as mais difíceis em termos de localização de bacilos pelo especialista humano. Isso acontece pois, embora a classificação das estruturas tenham apresentados valores semelhantes para imagens com muito e com poucos bacilos, a segmentação foi, em geral, melhor nas imagens com poucos bacilos.

6.1 Sugestões para trabalhos futuros

Como sugestões para trabalhos futuros, propõem-se os seguintes itens:

- A utilização de um banco de imagens diferentes, com imagens sem perda, haja visto que as imagens utilizadas nesse trabalho estavam em formato JPEG, que apresenta perdas de compressão e introdução de artefatos;
- O estudo de outras técnicas de segmentação, visando não eliminar as grandes áreas e, assim, trabalhar com bacilos isolados e grandes agrupamentos de bacilos; e

- O desenvolvimento de um método de identificação do melhor foco a ser utilizado, para o desenvolvimento de um sistema automático que não apenas seja capaz de identificar os bacilos de tuberculose na imagem, mas que também determine o foco adequado da câmera e o do microscópio quando da captura das imagens.

Referências

ALPAYDIN, E. *Introduction to machine learning*. [S.l.]: MIT press, 2004. Citado 4 vezes nas páginas 6, 33, 34 e 35.

CDC. *Tuberculosis (TB)*. 2014. Centers for Disease Control and Prevention - CDC. Disponível em <<http://www.cdc.gov/tb/>> [Acesso em 10 de setembro de 2014]. Citado na página 14.

CETELI. *An image database of conventional sputum smear microscopy for tuberculosis*. 2014. Center for Research and Development in Electronic and Information Technology - CETELI. Universidade Federal do Amazonas - UFAM. Disponível em <<http://www.tbimages.ufam.edu.br/>> [Acesso em 15 de abril de 2014]. Citado na página 37.

CHAYADEVI, M.; RAJU, G. Automated colour segmentation of tuberculosis bacteria thru region growing: A novel approach. In: IEEE. *Applications of Digital Information and Web Technologies (ICADIWT), 2014 Fifth International Conference on the*. [S.l.], 2014. p. 154–159. Citado 3 vezes nas páginas 20, 27 e 28.

COSTA FILHO, C. F. F. et al. Mycobacterium tuberculosis recognition with conventional microscopy. In: IEEE. *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. [S.l.], 2012. p. 6263–6268. Citado 9 vezes nas páginas 7, 15, 20, 27, 28, 37, 60, 61 e 62.

COSTA, L. F.; CESAR, R. M. *Shape analysis and classification: theory and practice*. [S.l.]: CRC press, 2000. Citado na página 30.

COSTA, M. G. et al. Automatic identification of mycobacterium tuberculosis with conventional light microscopy. In: IEEE. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. [S.l.], 2008. p. 382–385. Citado 7 vezes nas páginas 17, 20, 23, 37, 44, 60 e 61.

DESIKAN, P. Sputum smear microscopy in tuberculosis: Is it still relevant? In: . [S.l.: s.n.], 2013. v. 137, n. 13, p. 442–444. Citado na página 15.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. 2nd. ed. [S.l.]: Wiley, 2001. Citado 2 vezes nas páginas 23 e 24.

FAWCETT, T. An introduction to roc analysis. *Pattern recognition letters*, Elsevier, v. 27, n. 8, p. 861–874, 2006. Citado na página 48.

FORERO-VARGAS, M. G. et al. Automatic sputum color image segmentation for tuberculosis diagnosis. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *International Symposium on Optical Science and Technology*. [S.l.], 2001. p. 251–261. Citado na página 16.

FORERO-VARGAS, M. G. et al. Segmentation, autofocusing and signature extraction of tuberculosis sputum images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *International Symposium on Optical Science and Technology*. [S.l.], 2002. p. 171–182. Citado 5 vezes nas páginas 16, 20, 28, 29 e 62.

- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 3rd. ed. [S.l.]: Prentice Hall, 2007. Citado 6 vezes nas páginas 25, 26, 27, 28, 30 e 39.
- HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the marquardt algorithm. *Neural Networks, IEEE Transactions on*, IEEE, v. 5, n. 6, p. 989–993, 1994. Citado 2 vezes nas páginas 33 e 47.
- HAYKIN, S. S. et al. *Neural networks and learning machines*. [S.l.]: Pearson Education Upper Saddle River, 2009. Citado 3 vezes nas páginas 6, 31 e 32.
- HU, M.-K. Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, IEEE, v. 8, n. 2, p. 179–187, 1962. Citado na página 29.
- KHUTLANG, R. et al. Automated detection of tuberculosis in ziehl-neelsen-stained sputum smears using two one-class classifiers. *Journal of microscopy*, Wiley Online Library, v. 237, n. 1, p. 96–102, 2010. Citado 4 vezes nas páginas 17, 27, 28 e 29.
- KOSCHAN, A.; ABIDI, M. *Digital Color Image Processing*. [S.l.]: John Wiley and Sons, Inc., 2008. Citado na página 39.
- KUSWORO, A. et al. Autothresholding segmentation for tuberculosis bacteria identification in the ziehl-neelsen sputum sample. In: *Proceedings The 7th International Conference on Information & Communication Technology and Systems (ICTS)*. [S.l.: s.n.], 2013. p. 15–16. Citado 3 vezes nas páginas 20, 27 e 28.
- LEU, J.-G. Computing a shape's moments from its boundary. *Pattern recognition*, Elsevier, v. 24, n. 10, p. 949–957, 1991. Citado na página 28.
- MAKKAPATI, V.; AGRAWAL, R.; ACHARYA, R. Segmentation and classification of tuberculosis bacilli from zn-stained sputum smear images. In: IEEE. *Automation Science and Engineering, 2009. CASE 2009. IEEE International Conference on*. [S.l.], 2009. p. 217–220. Citado na página 18.
- MINISTÉRIO DA SAÚDE. *Manual Nacional de Vigilância Laboratorial da Tuberculose e Outras Micobactérias*. [S.l.]: Ministerio da Saude, 2008. Citado 2 vezes nas páginas 69 e 70.
- MINISTÉRIO DA SAÚDE. *Tuberculosis (TB)*. 2015. MS - Ministério da Saúde. Portal da Saúde. Disponível em <<http://portalsaude.saude.gov.br/>> [Acesso em 27 de fevereiro de 2015]. Citado na página 14.
- NAYAK, R.; SHENOY, V. P.; GALIGEKERE, R. R. A new algorithm for automatic assessment of the degree of tb-infection using images of zn-stained sputum smear. In: IEEE. *Systems in Medicine and Biology (ICSMB), 2010 International Conference on*. [S.l.], 2010. p. 294–299. Citado na página 18.
- OTSU, N. A threshold selection method from gray-level histograms. *Automatica*, v. 11, n. 285–296, p. 23–27, 1979. Citado na página 25.
- RULANINGTYAS, R.; SUKSMONO, A. B.; MENGKO, T. L. Automatic classification of tuberculosis bacteria using neural network. In: IEEE. *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*. [S.l.], 2011. p. 1–4. Citado 3 vezes nas páginas 19, 27 e 28.

- SADAPHAL, P. et al. Image processing techniques for identifying mycobacterium tuberculosis in ziehl-neelsen stains [short communication]. *The International Journal of Tuberculosis and Lung Disease*, International Union Against Tuberculosis and Lung Disease, v. 12, n. 5, p. 579–582, 2008. Citado 2 vezes nas páginas 17 e 27.
- SIENA, I. et al. Development of algorithm tuberculosis bacteria identification using color segmentation and neural networks. *International Journal of Video and Image Processing and Network Security*, v. 12, n. 4, p. 9–13, 2012. Citado 4 vezes nas páginas 19, 20, 27 e 28.
- SMART, T. *Background on smear microscopy in TB diagnosis*. 2007. Disponível em <<http://www.aidsmap.com/Background-on-smear-microscopy-in-TB-diagnosis/page/1426650/>> [Acesso em 27 de fevereiro de 2015]. Citado na página 15.
- SOTAQUIRÁ, M.; RUEDA, L.; NARVAEZ, R. Detection and quantification of bacilli and clusters present in sputum smear samples: a novel algorithm for pulmonary tuberculosis diagnosis. In: IEEE. *Digital Image Processing, 2009 International Conference on*. [S.l.], 2009. p. 117–121. Citado na página 18.
- STEWART, J. *Calculus. Early Transcendentals*. [S.l.]: Brooks/Cole, Cengage Learning, 2012. Citado 2 vezes nas páginas 6 e 28.
- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern recognition*. 4th. ed. [S.l.]: Elsevier, 2009. Citado na página 24.
- UNIVERSITY OF UTAH. *Acid-fast bacteria - Ziehl-Neelsen stain*. 2015. Health Sciences Library. Histology Procedure Manuals. Disponível em <<http://library.med.utah.edu/WebPath/HISTHTML/MANUALS/MANUALS.html>> [Acesso em 22 de abril de 2015]. Citado na página 15.
- VEROPOULOS, K. et al. The automated identification of tubercle bacilli using image processing and neural computing techniques. In: *ICANN 98*. [S.l.]: Springer, 1998. p. 797–802. Citado 3 vezes nas páginas 16, 29 e 62.
- WHO. *Global tuberculosis report 2014*. [S.l.]: World Health Organization, 2014. Citado na página 14.
- WHO. *Tuberculosis (TB)*. 2015. World Health Organization - WHO. Disponível em <<http://www.who.int/tb/>> [Acesso em 25 de fevereiro de 2015]. Citado 3 vezes nas páginas 4, 5 e 14.
- WITTEN, I. H.; FRANK, E. *Data Mining: Practical machine learning tools and techniques*. [S.l.]: Morgan Kaufmann, 2005. Citado na página 47.
- ZHAI, Y. et al. Automatic identification of mycobacterium tuberculosis from zn-stained sputum smear: Algorithm and system design. In: IEEE. *Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on*. [S.l.], 2010. p. 41–46. Citado 2 vezes nas páginas 19 e 28.

Apêndices

APÊNDICE A – Diagnóstico de tuberculose por baciloscopia

A.1 Preparação das lâminas

A baciloscopia é realizada, na maioria das vezes, com escarro espontâneo, utilizando a porção mais purulenta que é colocada e distendida diretamente sobre a lâmina. Isso evita resultados falsos negativos. Além disso, a utilização de lâminas sem uso prévio evita resultados falso-positivos (MINISTÉRIO DA SAÚDE, 2008).

O procedimento para a preparação das lâminas conforme o manual elaborado pelo Ministério da Saúde (2008) é descrito a seguir: primeiramente a partícula maior e mais purulenta da amostra de escarro é depositada e distendida na lâmina com movimento de vai e vem até obter um esfregaço homogêneo que cubra 2/3 da lâmina, sem deixar espaços vazios e, em seguida, é seca em temperatura ambiente.

Depois de secas, a lâmina deve ser fixada com o auxílio de um bico de Bunsen. A fixação é feita passando a lâmina, com o esfregaço voltado para cima, rapidamente, por três vezes, sobre a chama do bico de Bunsen. Depois de fixados, os esfregaços ficam aderidos às lâminas e estão prontos para a coloração.

A parede celular das micobactérias, tem em sua constituição química diversos lipídeos, entre eles os ácidos micólicos. A ligação de alguns desses lipídios com o corante Fucsina gera complexos que são responsáveis pela característica tintorial de resistência à descoloração por soluções álcool-ácidas, apresentada pelas células bacterianas que são então designadas como Bacilos Álcool-Ácido Resistentes (BAAR).

O método de Ziehl Neelsen baseia-se na propriedade de álcool-ácido resistência e sua utilização é recomendada pelo Ministério da Saúde para todos os laboratórios que realizam o diagnóstico da tuberculose pulmonar pelo Ministério da Saúde, uma vez que possibilita a identificação de BAAR e utiliza o microscópio ótico comum, para leitura do esfregaço.

Na coloração, todo o esfregaço presente na lâmina deve ser coberto com a Fucsina e, em seguida, uma chama deve ser passada lentamente por debaixo da lâmina até que ocorra a emissão de vapores visíveis. Após cinco minutos, a chama deve ser passada lentamente mais duas vezes por debaixo da lâmina. Após um minuto, deve-se derramar o excesso de Fucsina da lâmina e lavá-la com um pequeno filete de água corrente.

Em seguida, deve-se descolorir o esfregaço aplicando uma Solução Descorante de

Álcool-Ácido a 3% em toda a lâmina e novamente lavá-la com um pequeno filete de água corrente. Ao final, deve-se cobrir o esfregaço com o Azul de Metileno a 0,3% para colorir o fundo. Após 30 segundos, derrama-se o excesso de Azul de Metileno da lâmina e esta é lavada também com um pequeno filete de água corrente.

A.2 Leitura e interpretação dos resultados

Na leitura de todas as baciloscopias devem ser lidos no mínimo 100 campos úteis de microscópio, ou seja, aqueles campos nos quais se observam elementos celulares de origem pulmonar, de modo que os campos em que não aparecem esses elementos não devem ser contabilizados na leitura (MINISTÉRIO DA SAÚDE, 2008).

O resultado é interpretado da seguinte forma: quando não são encontrados BAAR em 100 campos, relata-se o resultado como negativo; quando são encontrados de 1 a 9 BAAR em 100 campos, relata-se apenas a quantidade de BAAR encontrada; quando são encontrados de 10 a 99 BAAR em 100 campos, relata-se o resultado como positivo +; quando é encontrado em média de 1 a 10 BAAR por campo, nos primeiros 50 campos observados, relata-se o resultado como positivo ++; e quando é encontrado em média mais de 10 BAAR por campo, nos primeiros 20 campos observados, relata-se o resultado como positivo +++. A leitura da lâmina deve ser feita da esquerda para a direita (MINISTÉRIO DA SAÚDE, 2008).

APÊNDICE B – Resultado da classificação das estruturas segmentadas

B.1 Classificação usando redes neurais

A Tabela 7 apresenta o resultado de sensibilidade obtido pela classificação das estruturas segmentadas pelo método baseado em limiarização global pela rede neural do tipo *perceptron* multicamadas com 1 camada oculta composta por 20 neurônios. A quantidade de neurônios na camada oculta foi escolhida através de testes exaustivos considerando diferentes topologias com relação ao número de neurônios. O valor de 20 neurônios na camada oculta representa o melhor compromisso entre resultados e esforço computacional causado pelo elevado número de neurônios.

As Tabelas 8, 9 e 10 apresentam os resultados de especificidade, acurácia e precisão, respectivamente, obtidos pela rede neural. A definição de cada um desses parâmetros foi apresentado na Seção 3.2.2.

As linhas representam o resultado de cada uma das 12 imagens para cada um dos 10 *folds*, ou seja, cada escolha diferente do conjunto de treino e do conjunto de teste, representados pelas colunas.

Para a sensibilidade, um resultado de 100% significa que todas as estruturas segmentadas pertencentes a bacilos foram classificadas como tais, enquanto um resultado de 0% significa que nenhuma estrutura pertencente a bacilo foi corretamente classificada. De forma análoga, para a especificidade, um resultado de 100% significa que todas as estruturas segmentadas não pertencentes a bacilos foram classificadas como tais, enquanto um resultado de 0% significa que nenhuma estrutura não pertencente a bacilo foi corretamente classificada. Para a acurácia, um resultado de 100% significa que todas as estruturas foram corretamente classificadas.

A precisão indica a porcentagem de estruturas corretamente classificadas como bacilos em relação a todas as estruturas classificadas como bacilos. O valor ausente na Tabela 10 (representado como um traço) é devido ao fato de que nessa imagem nenhuma estrutura foi classificada como bacilo. Uma precisão de 100% indica que todas as estruturas classificadas como bacilos são, de fato, bacilos.

Além dos resultados obtidos em cada imagem, as Tabelas 7, 8, 9 e 10 mostram os valores de média, máximo e mínimo para cada *fold* e as médias para cada um dos pacientes.

Tabela 7 – Sensibilidade devido à classificação por redes neurais.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	96,53%	92,47%	96,44%	96,77%	96,59%	94,16%	90,29%	95,65%	93,71%	81,30%
2	98,33%	100%	98,80%	92,11%	92,59%	77,27%	50,00%	93,48%	100%	98,21%
3	80,47%	96,77%	89,55%	84,73%	90,78%	91,95%	80,31%	95,74%	93,64%	93,94%
4	100%	75,00%	86,67%	60,00%	66,67%	78,57%	50,00%	72,73%	75,00%	66,67%
5	50,00%	50,00%	100%	100%	100%	100%	100%	100%	75,00%	100%
6	100%	64,52%	78,95%	60,00%	25,00%	0%	56,25%	46,15%	75,00%	36,36%
7	96,88%	100%	95,24%	100%	95,24%	93,75%	95,16%	96,67%	94,00%	95,12%
8	91,30%	86,67%	92,86%	100%	96,15%	93,33%	100%	100%	100%	79,31%
9	98,11%	63,79%	98,04%	91,84%	98,46%	100%	97,01%	98,25%	92,86%	96,43%
10	81,82%	80,00%	94,44%	92,31%	100%	93,75%	84,62%	75,00%	100%	100%
11	100%	100%	100%	100%	100%	75,00%	100%	50,00%	100%	100%
12	100%	50,00%	100%	80,00%	80,00%	100%	100%	100%	100%	66,67%
MÉDIA	91,12%	79,94%	94,25%	88,15%	86,79%	83,15%	83,64%	85,31%	91,60%	84,50%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	50,00%	50,00%	78,95%	60,00%	25,00%	0,00%	50,00%	46,15%	75,00%	36,36%

Tabela 8 – Especificidade devido à classificação por redes neurais.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	61,49%	58,54%	54,84%	58,82%	59,77%	64,55%	67,52%	87,50%	75,27%	77,17%
2	80,36%	83,33%	86,96%	84,95%	90,67%	91,00%	96,36%	89,66%	87,34%	82,29%
3	83,93%	71,88%	86,30%	92,77%	80,88%	75,00%	92,11%	77,42%	77,78%	79,41%
4	98,13%	99,43%	98,31%	89,80%	98,68%	98,85%	98,80%	99,56%	98,64%	99,50%
5	100%	99,02%	97,10%	100%	99,61%	93,42%	100%	100%	93,62%	99,36%
6	99,55%	99,46%	99,05%	99,38%	99,58%	100%	99,34%	98,17%	100%	99,17%
7	66,67%	100%	88,89%	100%	0%	80,00%	80,00%	75,00%	100%	100%
8	0%	100%	100%	66,67%	90,00%	100%	66,67%	100%	50,00%	100%
9	100%	100%	100%	96,00%	100%	100%	100%	100%	100%	100%
10	100%	100%	85,71%	0%	100%	0%	100%	100%	100%	100%
11	100%	100%	100%	100%	100%	95,65%	0%	100%	100%	99,61%
12	100%	100%	100%	100%	100%	100%	100%	100%	99,30%	100%
MÉDIA	82,51%	92,64%	91,43%	82,37%	84,93%	83,21%	83,40%	93,94%	90,16%	94,71%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	0,00%	58,54%	54,84%	0,00%	0,00%	0,00%	0,00%	75,00%	50,00%	77,17%

Tabela 9 – Acurácia devido à classificação por redes neurais.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	84,73%	82,09%	78,82%	86,25%	88,78%	86,36%	83,54%	94,12%	89,84%	79,89%
2	89,66%	89,84%	93,42%	87,02%	91,18%	85,54%	80,36%	91,35%	90,29%	88,16%
3	81,52%	90,40%	88,41%	87,85%	87,56%	90,30%	84,73%	91,20%	91,41%	90,96%
4	98,16%	98,88%	97,84%	87,04%	97,75%	97,34%	97,30%	98,33%	97,42%	99,25%
5	97,73%	98,08%	97,26%	100%	99,61%	93,51%	100%	100%	92,16%	99,37%
6	99,56%	96,74%	97,38%	98,19%	97,18%	94,12%	97,18%	96,19%	99,76%	97,78%
7	94,29%	100%	94,12%	100%	95,24%	91,89%	94,03%	92,11%	94,74%	95,35%
8	87,50%	93,65%	94,12%	97,14%	94,44%	94,29%	92,00%	100%	96,00%	81,82%
9	98,18%	70,42%	98,11%	93,24%	98,68%	100%	97,30%	98,44%	93,33%	96,72%
10	84,62%	88,00%	92,00%	92,31%	100%	93,75%	93,94%	78,26%	100%	100%
11	100%	100%	100%	100%	100%	92,59%	100%	98,88%	100%	99,61%
12	100%	99,66%	100%	98,88%	98,39%	100%	100%	100%	99,33%	98,21%
MÉDIA	92,99%	92,31%	94,29%	93,99%	95,73%	93,31%	93,36%	94,90%	95,36%	93,93%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	81,52%	70,42%	78,82%	86,25%	87,56%	85,54%	80,36%	78,26%	89,84%	79,89%

Tabela 10 – Precisão devido à classificação por redes neurais.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	83,15%	83,50%	74,39%	85,96%	89,91%	88,15%	86,85%	97,06%	93,45%	87,34%
2	84,29%	79,37%	90,11%	71,43%	78,13%	85,00%	87,88%	87,76%	70,59%	76,39%
3	91,96%	90,91%	92,31%	94,87%	90,78%	97,16%	94,44%	92,78%	96,26%	94,66%
4	50,00%	75,00%	68,42%	37,50%	60,00%	84,62%	57,14%	88,89%	75,00%	50,00%
5	100%	50,00%	66,67%	100%	50,00%	16,67%	100%	100%	50,00%	66,67%
6	71,43%	90,91%	88,24%	75,00%	66,67%	-	81,82%	50,00%	100%	50,00%
7	96,88%	100%	97,56%	100%	100%	96,77%	98,33%	93,55%	100%	100%
8	95,45%	100%	100%	96,97%	96,15%	100%	90,48%	100%	95,83%	100%
9	100%	100%	100%	97,83%	100%	100%	100%	100%	100%	100%
10	100%	100%	94,44%	100%	100%	100%	100%	100%	100%	100%
11	100%	100%	100%	100%	100%	75,00%	100%	100%	100%	50,00%
12	100%	100%	100%	100%	100%	100%	100%	100%	87,50%	100%
MÉDIA	89,43%	89,14%	89,34%	88,30%	85,97%	85,76%	91,41%	92,50%	89,05%	81,25%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	50,00%	50,00%	66,67%	37,50%	50,00%	16,67%	57,14%	50,00%	50,00%	50,00%

A Tabela 11 mostra a média e o desvio padrão da classificação das estruturas segmentadas para cada um dos 12 pacientes, onde foram destacados os piores valores para cada um dos índices avaliados.

Tabela 11 – Resumo dos valores de média e desvio padrão por paciente para a classificação das estruturas segmentadas usando redes neurais artificiais.

PACIENTE	SENSIBILIDADE		ESPECIFICIDADE		PRECISÃO		ACURÁCIA	
	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO
1	93,39%	4,75%	66,55%	10,36%	85,44%	4,68%	86,98%	6,15%
2	90,08%	15,63%	87,29%	4,80%	88,68%	3,69%	81,09%	6,95%
3	89,79%	6,00%	81,75%	6,98%	88,43%	3,21%	93,61%	2,20%
4	73,13%	13,88%	97,97%	2,91%	96,93%	3,54%	64,66%	16,53%
5	87,50%	21,25%	98,21%	2,63%	97,77%	2,81%	70,00%	29,19%
6	54,22%	28,79%	99,37%	0,52%	97,41%	1,63%	74,90%	17,44%
7	96,21%	2,22%	79,06%	30,32%	95,18%	2,79%	98,31%	2,16%
8	93,96%	6,89%	77,33%	32,80%	93,10%	5,13%	97,49%	3,16%
9	93,48%	10,74%	99,60%	1,26%	94,44%	8,73%	99,78%	0,69%
10	90,19%	9,19%	78,57%	41,65%	92,29%	7,13%	99,44%	1,76%
11	92,50%	16,87%	89,53%	31,49%	99,11%	2,32%	92,50%	16,87%
12	87,67%	17,92%	99,93%	0,22%	99,45%	0,71%	98,75%	3,95%

A Tabela 12 apresenta a taxa de classificação de estruturas duvidosas como bacilos, o que permite verificar o comportamento do sistema desenvolvido frente às estruturas duvidosas, segundo o especialista, que foram segmentadas.

Tabela 12 – Taxa de classificação de estruturas duvidosas como bacilos usando redes neurais.

	TAXA DE CLASSIFICAÇÃO
FOLD - 1	82,50%
FOLD - 2	45,22%
FOLD - 3	46,64%
FOLD - 4	49,84%
FOLD - 5	32,78%
FOLD - 6	48,48%
FOLD - 7	39,11%
FOLD - 8	64,52%
FOLD - 9	37,82%
FOLD - 10	43,10%

B.2 Classificação usando máquinas de vetores suporte

A Tabela 13 apresenta o resultado de sensibilidade obtido pela classificação das estruturas segmentadas pelo método baseado em limiarização global pela máquina de vetores suporte tendo como função de *kernel* uma função polinomial de grau 5. O *kernel* de função polinomial de grau 5 foi escolhido, após testes exaustivos, por apresentar, de

forma geral, melhores resultados de sensibilidade, especificidade, acurácia e precisão em comparação aos polinômios de grau inferior e superior e funções de base radial.

As Tabelas 14, 15 e 16 apresentam os resultados de especificidade, acurácia e precisão, respectivamente, obtidos pela máquina de vetor suporte descrita acima.

As linhas representam o resultado de cada uma das 12 imagens para cada um dos 10 *folds*, ou seja, cada escolha diferente do conjunto de treino e do conjunto de teste, representados pelas colunas.

A mesma análise feita para os resultados apresentados pela classificação usando redes neurais artificiais é válida para a classificação usando máquinas de vetores suporte.

Além dos resultados obtidos em cada imagem, as Tabelas 13, 14, 15 e 16 mostram os valores de média, máximo e mínimo para cada *fold* e as médias para cada um dos pacientes.

Tabela 13 – Sensibilidade devido à classificação por máquinas de vetores suporte.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	95,90%	94,98%	96,44%	96,14%	97,52%	94,48%	91,01%	94,57%	92,00%	79,67%
2	96,67%	98,00%	98,80%	89,47%	88,89%	75,76%	58,62%	100%	100%	98,21%
3	81,25%	95,74%	89,55%	86,26%	91,49%	93,29%	79,53%	96,81%	90,91%	95,45%
4	100%	75,00%	80,00%	60,00%	66,67%	64,29%	37,50%	72,73%	58,33%	66,67%
5	50,00%	50,00%	100%	100%	100%	100%	100%	66,67%	75,00%	100%
6	80,00%	54,84%	73,68%	60,00%	37,50%	0%	62,50%	46,15%	75,00%	45,45%
7	100%	100%	92,86%	100%	97,62%	96,88%	95,16%	96,67%	96,00%	97,56%
8	95,65%	80,00%	96,43%	93,75%	96,15%	93,33%	100%	100%	100%	86,21%
9	98,11%	65,52%	98,04%	89,80%	95,38%	100%	97,01%	100%	95,24%	98,21%
10	77,27%	86,67%	94,44%	92,31%	100%	100%	84,62%	65,00%	100%	100%
11	100%	100%	100%	100%	100%	75,00%	100%	50,00%	100%	100%
12	100%	50,00%	100%	80,00%	80,00%	100%	100%	100%	100%	66,67%
MÉDIA	89,57%	79,23%	93,35%	87,31%	87,60%	82,75%	83,83%	82,38%	90,21%	86,18%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	50,00%	50,00%	73,68%	60,00%	37,50%	0,00%	37,50%	46,15%	58,33%	45,45%

Tabela 14 – Especificidade devido à classificação por máquinas de vetores suporte.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	63,98%	58,54%	56,45%	59,66%	57,47%	61,82%	67,52%	82,81%	79,57%	77,17%
2	78,57%	82,05%	89,86%	84,04%	90,67%	94,00	96,36%	87,93%	87,34%	84,38%
3	83,93%	81,25%	86,30%	90,36%	80,88%	75,00%	90,79%	77,42%	72,22%	85,29%
4	97,50%	98,29%	97,47%	91,84%	98,34%	99,43%	99,20%	99,12%	98,19%	99,25%
5	100%	98,04%	92,75%	100%	99,61%	94,74%	100%	99,29%	95,74%	98,72%
6	99,78%	99,46%	98,57%	99,69%	99,58%	100%	99,34%	98,48%	100%	99,59%
7	100%	100%	88,89%	100%	0%	80,00%	80,00%	75,00%	100%	100%
8	0%	100%	100%	66,67%	90,00%	100%	66,67%	100%	50,00%	100%
9	100%	100%	100%	96,00%	100%	100%	100%	100%	100%	100%
10	100%	100%	85,71%	0%	100%	0%	100%	100%	100%	100%
11	100%	100%	100%	100%	100%	95,65%	0%	100%	100%	99,61%
12	100%	100%	100%	100%	100%	100%	100%	100%	99,30%	100%
MÉDIA	85,31%	93,13%	91,33%	82,36%	84,71%	83,39%	83,32%	93,34%	90,20%	95,33%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	0,00%	58,54%	56,45%	0,00%	0,00%	0,00%	0,00%	75,00%	50,00%	77,17%

Tabela 15 – Acurácia devido à classificação por máquinas de vetores suporte.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	85,15%	83,83%	79,50%	86,05%	89,02%	85,89%	84,05%	92,35%	89,39%	78,82%
2	87,93%	88,28%	94,74%	85,61%	90,20%	86,75%	83,33%	93,27%	90,29%	89,47%
3	82,07%	92,06%	88,41%	87,85%	88,04%	91,52%	83,74%	92,00%	88,28%	93,37%
4	97,55%	97,77%	96,77%	88,89%	97,43%	96,81%	97,30%	97,91%	96,14%	99,00%
5	97,73%	97,12%	93,15%	100%	99,61%	94,81%	100%	98,60%	94,12%	98,73%
6	99,56%	95,99%	96,51%	98,49%	97,58%	94,40%	97,49%	96,48%	99,76%	98,39%
7	100%	100%	92,16%	100%	97,62%	94,59%	94,03%	92,11%	96,49%	97,67%
8	91,67%	90,48%	97,06%	91,43%	94,44%	94,29%	92,00%	100%	96,00%	87,88%
9	98,18%	71,83%	98,11%	91,89%	96,05%	100%	97,30%	100%	95,56%	98,36%
10	80,77%	92,00%	92,00%	92,31%	100%	100%	93,94%	69,57%	100%	100%
11	100%	100%	100%	100%	100%	92,59%	100%	98,88%	100%	99,61%
12	100%	99,66%	100%	98,88%	98,39%	100%	100%	100%	99,33%	98,21%
MÉDIA	93,38%	92,42%	94,03%	93,45%	95,70%	94,30%	93,60%	94,26%	95,45%	94,96%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	80,77%	71,83%	79,50%	85,61%	88,04%	85,89%	83,33%	69,57%	88,28%	78,82%

Tabela 16 – Precisão devido à classificação por máquinas de vetores suporte.

PACIENTE	<i>FOLD</i>									
	1	2	3	4	5	6	7	8	9	10
1	83,98%	83,86%	75,08%	86,17%	89,49%	87,39%	86,94%	95,96%	94,43%	87,11%
2	82,86%	77,78%	92,13%	69,39%	77,42%	89,29%	89,47%	86,79%	70,59%	78,57%
3	92,04%	93,75%	92,31%	93,39%	90,85%	97,20%	93,52%	92,86%	95,24%	96,18%
4	42,86%	50,00%	57,14%	42,86%	54,55%	90,00%	60,00%	80,00%	63,64%	40,00%
5	100%	33,33%	44,44%	100%	50,00%	20,00%	100%	66,67%	60,00%	50,00%
6	80,00%	89,47%	82,35%	85,71%	75,00%	-	83,33%	54,55%	100%	71,43%
7	100%	100%	97,50%	100%	100%	96,88%	98,33%	93,55%	100%	100%
8	95,65%	100%	100%	96,77%	96,15%	100%	90,48%	100%	95,83%	100%
9	100%	100%	100%	97,78%	100%	100%	100%	100%	100%	100%
10	100%	100%	94,44%	100%	100%	100%	100%	100%	100%	100%
11	100%	100%	100%	100%	100%	75,00%	100%	100%	100%	50,00%
12	100%	100%	100%	100%	100%	100%	100%	100%	87,50%	100%
MÉDIA	89,78%	85,68%	86,28%	89,34%	86,12%	86,89%	91,84%	89,20%	88,94%	81,11%
MÁXIMO	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
MÍNIMO	42,86%	33,33%	44,44%	42,86%	50,00%	20,00%	60,00%	54,55%	60,00%	40,00%

A Tabela 17 mostra a média e o desvio padrão da classificação das estruturas segmentadas para cada um dos 12 pacientes, onde foram destacados os piores valores para cada um dos índices avaliados.

Tabela 17 – Resumo dos valores de média e desvio padrão por paciente para a classificação das estruturas segmentadas usando máquinas de vetores suporte.

PACIENTE	SENSIBILIDADE		ESPECIFICIDADE		PRECISÃO		ACURÁCIA	
	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO	MÉDIA	DESVIO PADRÃO
1	93,27%	5,17%	66,50%	9,85%	85,40%	4,22%	87,04%	5,81%
2	90,44%	13,50%	87,52%	5,45%	88,99%	3,41%	81,43%	7,97%
3	90,03%	5,99%	82,34%	6,20%	88,73%	3,68%	93,73%	1,96%
4	68,12%	16,14%	97,86%	2,23%	96,55%	2,80%	58,10%	16,34%
5	84,17%	21,68%	97,89%	2,58%	97,39%	2,53%	62,44%	28,94%
6	53,51%	23,41%	99,45%	0,53%	97,47%	1,67%	80,21%	12,67%
7	97,27%	2,32%	82,39%	30,64%	96,47%	3,11%	98,63%	2,15%
8	94,15%	6,48%	77,33%	32,80%	93,52%	3,55%	97,49%	3,15%
9	93,73%	10,35%	99,60%	1,26%	94,73%	8,39%	99,78%	0,70%
10	90,03%	11,80%	78,57%	41,65%	92,06%	9,94%	99,44%	1,76%
11	92,50%	16,87%	89,53%	31,49%	99,11%	2,32%	92,50%	16,87%
12	87,67%	17,92%	99,93%	0,22%	99,45%	0,71%	98,75%	3,95%

A Tabela 18 apresenta a taxa de classificação de estruturas duvidosas como bacilos, o que permite verificar o comportamento do sistema desenvolvido frente às estruturas duvidosas, segundo o especialista, que foram segmentadas.

Tabela 18 – Taxa de classificação de estruturas duvidosas como bacilos usando máquinas de vetores suporte.

	TAXA DE CLASSIFICAÇÃO
FOLD - 1	78,75%
FOLD - 2	45,37%
FOLD - 3	47,89%
FOLD - 4	48,95%
FOLD - 5	36,53%
FOLD - 6	52,15%
FOLD - 7	35,87%
FOLD - 8	70,24%
FOLD - 9	28,16%
FOLD - 10	41,67%